(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 March 2002 (14.03.2002)

PCT

(10) International Publication Number
WO 02/20731 A2

(51) International Patent Classification[7]: C12N 9/00

(21) International Application Number: PCT/EP01/10125

(22) International Filing Date: 3 September 2001 (03.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/231,297 8 September 2000 (08.09.2000) US

(71) Applicant *(for all designated States except US)*: BAYER AKTIENGESELLSCHAFT [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: XIAO, Yonghong [US/US]; 78 Dana Street #1, Cambridge, MA 02138 (US).

(74) Common Representative: BAYER AKTIENGESELLSCHAFT; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *entirely in electronic form (except for this front page) and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/20731 A2

(54) Title: REGULATION OF HUMAN PROTEIN DISULFIDE ISOMERASE-LIKE ENZYME

(57) Abstract: Reagents which regulate human protein disulfide isomerase-like enzyme and reagents which bind to human protein disulfide isomerase-like enzyme gene products can play a role in preventing, ameliorating, or correcting dysfunctions or disorders including, but not limited to, cancer, diabetes, and obesity.

REGULATION OF HUMAN PROTEIN DISULFIDE ISOMERASE-LIKE ENZYME

TECHNICAL FIELD OF THE INVENTION

The invention relates to the area of regulation of human protein disulfide isomerase-like enzyme to provide therapeutic effects.

BACKGROUND OF THE INVENTION

Cells contain a number of proteins known as molecular chaperones or foldases. U.S. Patent 6,001,632. These molecules catalyze the folding of newly synthesized proteins, prevent aggregation and improper glycosylation, and remove denatured proteins. Although they do not become part of the final structure, they are important in the assembly of proteins or their subunits into larger, more complex structures. In the absence of chaperones and foldases, misfolded proteins are quickly degraded by intracellular proteases.

Foldases, such as protein disulfide isomerase (PDI), are specialized enzymes which carry out rate-limiting covalent steps in protein folding. These enzymes are most abundant in cells actively synthesizing secreted proteins which are major components of the ER lumen. Tasanen *et al.*, *J. Biol. Chem. 267*, 11513-19, 1992) and may constitute 1-2% of eukaryotic cellular proteins. Although incubation of reduced unfolded proteins in buffers with defined ratios of oxidized and reduced thiols can lead to native conformation, the rate of folding is slow and the attainment of native conformation decreases proportionately to the size and number of cysteines in the protein. In contrast, PDI in the eukaryotic ER is much more efficient in carrying out the enzymatic pairing and oxidation of cysteines.

Human PDI is a resident endoplasmic reticulum protein which mediates the formation of disulfide bonds in recently synthesized proteins. Chen *et al., Br. J. Haematol. 90,* 425-31, 1995. PDI has recently been localized additional sites. Catalytically active PDI also has been localized to the plasma membrane of mammalian cells and platelets. PDI is also released from activated and inactivated platelets. *Id.* PDI catalyzes the reversible formation and isomerization of disulfide bonds, thus promoting correct protein folding. *Id.*; Essex *et al., Blood 86,* 2168-73,1995.

The activity of PDI is extremely concentration-dependent under certain conditions. PDI can promote the misfolding and aggregation of its substrates. When PDI is present at high concentrations relative to substrate, it acts a chaperone, increasing the yield of properly folded protein, and inhibiting substrate aggregation. In contrast when PDI is present at substoichiometric concentrations relative to substrate, it facilitates aggregation and inhibits proper protein folding. Sideraki & Gilbert, *Biochem. 39,* 1180-88, 2000.

PDI plays an important role in platelet responses. PDI has been shown to mediate platelet aggregation and secretion. PDI is further postulated to promote wound healing by facilitating the proper folding of proteins important for tissue remodeling. Essex & Li, *Br. J. Haematol. 104,* 448-54, 1999.

PDI also has been implicated tumor cell proliferation. Tumor cells responsive to the potent antiproliferative Tissue Growth Factor β1 (TGFβ1) also down regulate PDI expression when treated with TGFβ1. TGFβ1-resistant tumor cells, however do not alter their pattern of PDI expression after treatment with TGFβ1. PDI is postulated to interfere with TGFβ1 activity by either altering the folding of the TGFβ1 receptor or the TGFβ1 molecule itself. Sipes *et al., Cell Growth & Diff. 1,* 241-46, 1990.

Because of the importance of PDI in a variety of cellular processes, there is a need in the art to identify additional related molecules which can be regulated to provide therapeutic effects.

SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human protein disulfide isomerase-like enzyme. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a protein disulfide isomerase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2,

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29,

the amino acid sequence shown in SEQ ID NO:2; and

the amino acid sequence shown in SEQ ID NO:29.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a protein disulfide isomerase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2,

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29,

the amino acid sequence shown in SEQ ID NO:2; and

the amino acid sequence shown in SEQ ID NO:29.

Binding between the test compound and the protein disulfide isomerase-like enzyme polypeptide is detected. A test compound which binds to the protein disulfide isomerase-like enzyme polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the protein disulfide isomerase-like enzyme.

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO:1,

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO:28,

the nucleotide sequence shown in SEQ ID NO:1; and

the nucleotide sequence shown in SEQ ID NO:28.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the protein disulfide isomerase-like enzyme through interacting with the protein disulfide isomerase-like enzyme mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a protein disulfide isomerase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2,

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29,

the amino acid sequence shown in SEQ ID NO:2; and

the amino acid sequence shown in SEQ ID NO:29.

A protein disulfide isomerase-like enzyme activity of the polypeptide is detected. A test compound which increases protein disulfide isomerase-like enzyme activity of the polypeptide relative to protein disulfide isomerase-like enzyme activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases protein disulfide isomerase-like enzyme activity of the polypeptide relative to protein disulfide isomerase-like enzyme activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a protein disulfide isomerase-like enzyme product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO:1,

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO:28,

the nucleotide sequence shown in SEQ ID NO:1; and

the nucleotide sequence shown in SEQ ID NO:28.

Binding of the test compound to the protein disulfide isomerase-like enzyme product is detected. A test compound which binds to the protein disulfide isomerase-like enzyme product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO:1,

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO:28,

the nucleotide sequence shown in SEQ ID NO:1; and

the nucleotide sequence shown in SEQ ID NO:28.

Protein disulfide isomerase-like enzyme activity in the cell is thereby decreased.

The invention thus provides a human protein disulfide isomerase-like enzyme which can be used to identify test compounds which may act, for example, as agonists or antagonists at the enzyme's active site. Human protein disulfide isomerase-like enzyme and fragments thereof also are useful in raising specific antibodies which can block the enzyme and effectively reduce its activity.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:1).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO:2).

Fig. 3 shows the amino acid sequence of the protein identified by SwissProt Accession No. Q10057 (SEQ ID NO:3).

Fig. 4 shows the amino acid sequence of pfam/hmm/thiored (SEQ ID NO:4).

Fig. 5 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:5).

Fig. 6 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:6).

Fig. 7 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:7).

Fig. 8 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:8).

Fig. 9 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:9).

Fig. 10 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:10).

Fig. 11 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:11).

Fig. 12 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:12).

Fig. 13 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:13).

Fig. 14 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:14).

Fig. 15 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:15).

Fig. 16 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:16).

Fig. 17 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:17).

Fig. 18 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:18).

Fig. 19 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:19).

Fig. 20 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:20).

Fig. 21 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:21).

Fig. 22 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:22).

Fig. 23 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:23).

Fig. 24 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:24).

Fig. 25 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:25).

Fig. 26 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:26).

Fig. 27 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:27).

Fig. 28 shows the BLASTP alignment of 59_PROTEIN (SEQ ID NO:2) against swiss|Q10057|PDI1_SCHPO (SEQ ID NO:3).

Fig. 29 shows the BLOCKS search results.

Fig. 30 shows the HMMPFAM - alignment of 59_PROTEIN (SEQ ID NO:2) against pfam|hmm|thiored (SEQ ID NO:4)

Fig. 31 shows the DNA-sequence encoding a protein disulfide isomerase-like enzyme polypeptide (SEQ ID NO:28).

Fig. 32 shows the amino acid sequence deduced from the DNA-sequence of Fig. 31 (SEQ ID NO:29).

Fig. 33 shows the BLASTP alignment of 59_short_TR1 against swiss/P30101/ER60_HUMAN.

Fig. 34 shows the HMMPFAM alignment of 59_short_TR1 against pfam/hmm/thiored.

Fig. 35 shows the expression profile of protein disulfide isomerase-like protein mRNA in whole body tissues.

Fig. 36 shows the expression profile of protein disulfide isomerase-like protein mRNA in blood and lung tissues.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

   amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2,
   amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29,

      the amino acid sequence shown in SEQ ID NO:2; and
      the amino acid sequence shown in SEQ ID NO:29.

b) a polynucleotide comprising the sequence of SEQ ID NO:1 or SEQ ID NO:28;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that a novel protein disulfide isomerase-like enzyme, particularly a human protein disulfide isomerase-like enzyme, is a discovery of the present invention. Human protein disulfide isomerase-like enzyme comprises the amino acid sequences shown in SEQ ID NO:2 and SEQ ID NO:29, as encoded by the coding sequence shown in SEQ ID NO:1 and SEQ ID NO:28, respectively.

Human protein disulfide isomerase-like enzyme is 32% identical over a 131 amino acid overlap to the protein identified by SwissProt Accession No. Q10057 (SEQ ID NO:3) and annotated as "PUTATIVE PROTEIN DISULFIDE ISOMERASE C1F5.02 PRECURSOR (EC 5.3.4.1)." (FIG. 28).

The human protein disulfide isomerase-like enzyme of the invention is expected to be useful for the same purposes as previously identified protein disulfide isomerase enzymes. For example, protein disulfide isomerase is found on the surface of platelets and is released upon platelet activation and mediates platelet aggregation and secretion. Because platelets are present and become activated when and where vascular injury occurs, platelet protein disulfide isomerase and, therefore, human protein disulfide isomerase-like enzyme, may play a role in the various hemostatic and tissue remodeling processes in which platelets (hence blood coagulation) are involved. Protein disulfide isomerase also has a regulatory effect on the surface protein thiol status of B cells. The increased expression of protein disulfide isomerase may play a crucial role in SH-mediated protection and drug resistance in malignant B lymphocytes. Protein disulfide isomerase mRNA levels correlate inversely with sensitivity to the growth-inhibitory effects of TGFβ 1, and protein disulfide isomerase down-regulation occurs in some TGF β-sensitive cells, but not in a human pancreatic carcinoma cell line which is insensitive to the growth-inhibitory effects of TGF β 1.

Because human protein disulfide isomerase-like enzyme is expressed mainly in tumor and other highly proliferative tissues (*i.e.*, tissues with a rich blood supply), it

is a potential target for diseases therapeutic sensitivity/resistance involving blood clotting, vascular wall repair and formation, as well as tissue remodeling. Human protein disulfide isomerase-like enzyme therefore can be used in therapeutic methods to treat disorders such as hematological disorders, cardiovascular disorders, and cancer. Human protein disulfide isomerase-like enzyme also can be used to screen for human protein disulfide isomerase-like enzyme agonists and antagonists.

Polypeptides

Protein disulfide isomerase-like polypeptides according to the invention comprise at least 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, or 390 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29 or a biologically active variant thereof, as defined below. A protein disulfide isomerase-like polypeptide of the invention therefore can be a portion of a protein disulfide isomerase-like enzyme, a full-length protein disulfide isomerase-like enzyme, or a fusion protein comprising all or a portion of a protein disulfide isomerase-like enzyme.

Biologically Active Variants

Protein disulfide isomerase-like polypeptide variants which are biologically active, *i.e.*, retain a protein disulfide isomerase-like activity, also are protein disulfide isomerase-like polypeptides. Preferably, naturally or non-naturally occurring protein disulfide isomerase-like polypeptide variants have amino acid sequences which are at least about 50, 55, 60, 65, or 70, preferably about 75, 80, 85, 90, 96, 96, or 98% identical to the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29 or a fragment thereof. Percent identity between a putative protein disulfide isomerase-like polypeptide variant and an amino acid sequence of SEQ ID NO:2 or SEQ ID NO:29 is determined using the Blast2 alignment program (Blosum62, Expect 10, standard genetic codes).

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a protein disulfide isomerase-like polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active protein disulfide isomerase-like polypeptide can readily be determined by assaying for protein disulfide isomerase-like activity, as described for example, in Sideraki & Gilbert, 2000, Essex *et al.*, 1995, or Chen *et al.*, 1992.

Fusion Proteins

Fusion proteins are useful for generating antibodies against protein disulfide isomerase-like polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of a protein disulfide isomerase-like polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A protein disulfide isomerase-like polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, or 390 contiguous amino acids of SEQ ID NO:2 or SEQ ID NO:29 or

of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length protein disulfide isomerase-like enzyme protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horse-radish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the protein disulfide isomerase-like polypeptide-encoding sequence and the heterologous protein sequence, so that the protein disulfide isomerase-like polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO:1 or SEQ ID NO:28 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human protein disulfide isomerase-like polypeptide can be obtained using protein disulfide isomerase-like polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of protein disulfide isomerase-like polypeptide, and expressing the cDNAs as is known in the art.

Polynucleotides

A protein disulfide isomerase-like polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a protein disulfide isomerase-like polypeptide. A coding sequence for human protein disulfide isomerase-like enzyme is shown in SEQ ID NO:1 and SEQ ID NO:28. This coding sequence is found within the larger genomic sequence shown in SEQ ID NO:5.

Degenerate nucleotide sequences encoding human protein disulfide isomerase-like polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:28 also are protein disulfide isomerase-like polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of protein disulfide isomerase-like polynucleotides which encode biologically active protein disulfide isomerase-like polypeptides also are protein disulfide isomerase-like polynucleotides.

Identification of Polynucleotide Variants and Homologs

Variants and homologs of the protein disulfide isomerase-like polynucleotides described above also are protein disulfide isomerase-like polynucleotides. Typically, homologous protein disulfide isomerase-like polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known protein disulfide isomerase-like polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions: 2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 EC once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the protein disulfide isomerase-like polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of protein disulfide isomerase-like polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology (Bonner *et al., J. Mol. Biol. 81*, 123 (1973). Variants of human protein disulfide isomerase-like polynucleotides or protein disulfide isomerase-like polynucleotides of other species can therefore be identified by hybridizing a putative homologous protein disulfide isomerase-like polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:28 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to protein disulfide isomerase-like polynucleotides or their complements following stringent hybridization and/or wash conditions also are protein disulfide isomerase-like polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a protein disulfide isomerase-like polynucleotide having a nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:28 or the complement thereof and a polynucleotide sequence which is at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

$$T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%formamide) - 600/l),$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50% formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

<u>Preparation of Polynucleotides</u>

A naturally occurring protein disulfide isomerase-like polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such

technique for obtaining a polynucleotide can be used to obtain isolated protein disulfide isomerase-like polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments which comprises protein disulfide isomerase-like enzyme nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

Protein disulfide isomerase-like enzyme cDNA molecules can be made with standard molecular biology techniques, using protein disulfide isomerase-like enzyme mRNA as a template. Protein disulfide isomerase-like enzyme cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesizes protein disulfide isomerase-like polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a protein disulfide isomerase-like polypeptide having, for example, an amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29 or a biologically active variant thereof.

<u>Extending Polynucleotides</u>

The partial nucleotide sequence disclosed herein can be used to identify the corresponding full length gene from which they were derived. The partial sequence can be nick-translated or end-labeled with ^{32}P using polynucleotide kinase using labeling methods known to those with skill in the art (BASIC METHODS IN MOLECULAR BIOLOGY, Davis *et al.*, eds., Elsevier Press, N.Y., 1986). A lambda library prepared from human tissue can be directly screened with the labeled sequences of interest or the library can be converted en masse to pBluescript (Stratagene Cloning Systems, La Jolla, Calif. 92037) to facilitate bacterial colony

screening (see Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Press (1989, pg. 1.20).

Both methods are well known in the art. Briefly, filters with bacterial colonies containing the library in pBluescript or bacterial lawns containing lambda plaques are denatured, and the DNA is fixed to the filters. The filters are hybridized with the labeled probe using hybridization conditions described by Davis *et al.*, 1986. The partial sequences, cloned into lambda or pBluescript, can be used as positive controls to assess background binding and to adjust the hybridization and washing stringencies necessary for accurate clone identification. The resulting autoradiograms are compared to duplicate plates of colonies or plaques; each exposed spot corresponds to a positive colony or plaque. The colonies or plaques are selected, expanded and the DNA is isolated from the colonies for further analysis and sequencing.

Positive cDNA clones are analyzed to determine the amount of additional sequence they contain using PCR with one primer from the partial sequence and the other primer from the vector. Clones with a larger vector-insert PCR product than the original partial sequence are analyzed by restriction digestion and DNA sequencing to determine whether they contain an insert of the same size or similar as the mRNA size determined from Northern blot Analysis.

Once one or more overlapping cDNA clones are identified, the complete sequence of the clones can be determined , for example after exonuclease III digestion (McCombie *et al.*, *Methods 3*, 33-40, 1991). A series of deletion clones are generated, each of which is sequenced. The resulting overlapping sequences are assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a highly accurate final sequence.

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 2230 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 30553060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to

walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' nontranscribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Obtaining Polypeptides

Protein disulfide isomerase-like polypeptides can be obtained, for example, by purification from human cells, by expression of protein disulfide isomerase-like polynucleotides, or by direct chemical synthesis.

Protein Purification

Protein disulfide isomerase-like polypeptides can be purified from any cell which expresses the enzyme, including host cells which have been transfected with protein disulfide isomerase-like enzyme expression constructs. Testis, colon tumor, lung tumor, neuroendocrine lung carcinoid, glioblastoma, immortalized keratinocytes, normal lung, fetal liver, spleen, pancreatic adenocarcinoma, normal prostate, germinal center B cells, germ cell tumors, papillary serous carcinoma, breast tumor, kidney tumors, infant brain cDNA from a muscular atrophy patient, multiple sclerosis lesions, fetal brain, uterus, Burkitt's lymphoma, and choriocarcinoma provide useful sources of protein disulfide isomerase-like enzyme. A purified protein disulfide isomerase-like polypeptide is separated from other compounds which normally associate with the protein disulfide isomerase-like polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified protein disulfide isomerase-like polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of Polynucleotides

To express a protein disulfide isomerase-like polynucleotide, the polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding protein disulfide isomerase-like polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et*

al. (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a protein disulfide isomerase-like polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those nontranslated regions of the vector enhancers, promoters, 5' and 3' untranslated regions which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a protein disulfide isomerase-like polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the protein disulfide isomerase-like polypeptide. For example, when a large quantity of a protein disulfide isomerase-like polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the protein disulfide isomerase-like polypeptide can be ligated into the vector in frame with sequences for the amino terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 55035509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516544, 1987.

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding protein disulfide isomerase-like polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be

used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 16711680, 1984; Broglie *et al.*, *Science 224*, 838843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in McGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191196, 1992).

An insect system also can be used to express a protein disulfide isomerase-like polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding protein disulfide isomerase-like polypeptides can be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of protein disulfide isomerase-like polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which protein disulfide isomerase-like polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 32243227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be used to express protein disulfide isomerase-like polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding protein disulfide isomerase-like polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a protein disulfide isomerase-like polypeptide in

infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci. 81*, 36553659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding protein disulfide isomerase-like polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a protein disulfide isomerase-like polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ. 20*, 125162, 1994).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein disulfide isomerase-like polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Posttranslational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or

function. Different host cells which have specific cellular machinery and characteristic mechanisms for Post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express protein disulfide isomerase-like polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 12 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced protein disulfide isomerase-like enzyme sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 22332, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 81723, 1980) genes which can be employed in *tk*⁻ or *aprt*⁻ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 356770, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 114, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of

tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 804751, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol. 55*, 121131, 1995).

Detecting Expression

Although the presence of marker gene expression suggests that the protein disulfide isomerase-like polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a protein disulfide isomerase-like polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode a protein disulfide isomerase-like polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a protein disulfide isomerase-like polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the protein disulfide isomerase-like polynucleotide.

Alternatively, host cells which contain a protein disulfide isomerase-like polynucleotide and which express a protein disulfide isomerase-like polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a protein disulfide isomerase-like polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a protein disulfide isomerase-like polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from

sequences encoding a protein disulfide isomerase-like polypeptide to detect transformants which contain a protein disulfide isomerase-like polynucleotide.

A variety of protocols for detecting and measuring the expression of a protein disulfide isomerase-like polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a protein disulfide isomerase-like polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 12111216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding protein disulfide isomerase-like polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a protein disulfide isomerase-like polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a protein disulfide isomerase-like polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode protein disulfide isomerase-like polypeptides can be designed to contain signal sequences which direct secretion of soluble protein disulfide isomerase-like polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound protein disulfide isomerase-like polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a protein disulfide isomerase-like polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the protein disulfide isomerase-like polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a protein disulfide isomerase-like polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif. 3*, 263281, 1992), while the enterokinase cleavage site provides a means for purifying the protein disulfide isomerase-like polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol. 12*, 441453, 1993.

Chemical Synthesis

Sequences encoding a protein disulfide isomerase-like polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225232, 1980). Alternatively, a protein disulfide isomerase-like polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 21492154, 1963; Roberge *et al.*, *Science 269*, 202204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of protein disulfide isomerase-like polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic protein disulfide isomerase-like polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the protein disulfide isomerase-like polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce protein disulfide isomerase-like polypeptide-encoding nucleotide sequences possessing non-natural occurring codons. For example, codons preferred by a

particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter protein disulfide isomerase-like polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a protein disulfide isomerase-like polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a protein disulfide isomerase-like polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a protein disulfide isomerase-like polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well

known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a protein disulfide isomerase-like polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to protein disulfide isomerase-like polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a protein disulfide isomerase-like polypeptide from solution.

Protein disulfide isomerase-like polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a protein disulfide isomerase-like polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies which specifically bind to a protein disulfide isomerase-like polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B cell hybridoma technique, and the EBV hybridoma technique (Kohler *et al.*, *Nature 256*, 495497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 3142, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 20262030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 68516855, 1984; Neuberger *et al.*, *Nature 312*, 604608, 1984; Takeda *et al.*, *Nature 314*, 452454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies which specifically bind to a protein disulfide isomerase-like polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. Patent 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to protein disulfide isomerase-like polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 1112023, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91).

Antibodies which specifically bind to protein disulfide isomerase-like polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 38333837, 1989; Winter *et al.*, *Nature 349*, 293299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a protein disulfide isomerase-like polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes

and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of protein disulfide isomerase-like enzyme gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 18, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543583, 1990.

Modifications of protein disulfide isomerase-like enzyme gene expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the protein disulfide isomerase-like enzyme gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions 10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a protein disulfide isomerase-like polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a protein disulfide isomerase-like polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent protein disulfide isomerase-like enzyme nucleotides, can provide sufficient targeting specificity for protein disulfide isomerase-like enzyme mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular protein disulfide isomerase-like polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a protein disulfide isomerase-like polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5' substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 35393542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 15321539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a protein disulfide isomerase-like polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the protein disulfide isomerase-like polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a protein disulfide isomerase-like enzyme RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate protein disulfide isomerase-like enzyme RNA targets also can be evaluated

by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease protein disulfide isomerase-like enzyme expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human protein disulfide isomerase-like enzyme. Such genes may represent genes which are differentially expressed in disorders including, but not limited to, cancer, diabetes, and obesity. Further, such genes may represent genes which are

differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human protein disulfide isomerase-like gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed.,, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and,

preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human protein disulfide isomerase-like enzyme. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human protein disulfide isomerase-like enzyme. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human protein disulfide isomerase-like gene or gene product are up-regulated or down-regulated.

Screening Methods

The invention provides assays for screening test compounds which bind to or modulate the activity of a protein disulfide isomerase-like polypeptide or a protein disulfide isomerase-like polynucleotide. A test compound preferably binds to a protein disulfide isomerase-like polypeptide or polynucleotide. More preferably, a test compound decreases or increases protein disulfide isomerase-like activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially

addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution (*see, e.g.*, Houghten, *BioTechniques 13*, 412421, 1992), or on beads (Lam, *Nature 354*, 8284, 1991), chips (Fodor, *Nature 364*, 555556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 18651869, 1992), or phage (Scott & Smith, *Science 249*, 386390, 1990; Devlin, *Science 249*, 404406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 63786382, 1990; Felici, *J. Mol. Biol. 222*, 301310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to protein disulfide isomerase-like polypeptides or polynucleotides or to affect protein disulfide isomerase-like enzyme activity or protein disulfide isomerase-like enzyme gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials,

pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19,* 161418 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 710, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 5763 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support.

When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupies, for example, the ATP/GTP binding site of the enzyme or the active site of the protein disulfide isomerase-like polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the protein disulfide isomerase-like polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to the protein disulfide isomerase-like polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a protein disulfide isomerase-like polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a protein disulfide isomerase-like polypeptide. A microphysiometer (*e.g.*, CytosensorJ) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a protein disulfide isomerase-like polypeptide (McConnell *et al.*, *Science 257*, 19061912, 1992).

Determining the ability of a test compound to bind to a protein disulfide isomerase-like polypeptide also can be accomplished using a technology such as real-time

Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 23382345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a protein disulfide isomerase-like polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 1204612054, 1993; Bartel *et al.*, *BioTechniques 14*, 920924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 16931696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the protein disulfide isomerase-like polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a protein disulfide isomerase-like polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-BINDING and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the protein disulfide isomerase-like polypeptide.

It may be desirable to immobilize either the protein disulfide isomerase-like polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the protein disulfide isomerase-like polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the protein disulfide isomerase-like polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a protein disulfide isomerase-like polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the protein disulfide isomerase-like polypeptide is a fusion protein comprising a domain that allows the protein disulfide isomerase-like polypeptide to be bound to a solid support. For example, glutathione S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the nonadsorbed protein disulfide isomerase-like polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a protein disulfide isomerase-like polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated protein disulfide isomerase-like polypeptides (or polynucleotides) or test compounds can be prepared from biotinNHS(Nhydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a protein disulfide isomerase-like polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the ATP/GTP binding site or the active site of the protein disulfide isomerase-like polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the protein disulfide isomerase-like polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the protein disulfide isomerase-like polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a protein disulfide isomerase-like polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a protein disulfide isomerase-like polypeptide or polynucleotide can be used in a cell-based assay system. A protein disulfide isomerase-like polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a protein disulfide isomerase-like polypeptide or polynucleotide is determined as described above.

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease the protein disulfide isomerase-like activity of a human protein disulfide isomerase-like polypeptide. Protein disulfide isomerase-like activity can be measured, for example, as described in the specific examples, below, and in Sideraki & Gilbert, 2000, Essex *et al.*, 1995, or Chen *et al.*, 1992. Enzyme assays can be carried out after contacting either a purified protein disulfide isomerase-like polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test compound which decreases a transketolase activity of a protein disulfide isomerase-like polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing protein disulfide isomerase-like enzyme activity. A test compound which increases a transketolase activity of a human protein disulfide isomerase-like polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing human protein disulfide isomerase-like enzyme activity.

Gene Expression

In another embodiment, test compounds which increase or decrease protein disulfide isomerase-like enzyme gene expression are identified. A protein disulfide isomerase-like polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the protein disulfide isomerase-like polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of

the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of protein disulfide isomerase-like enzyme mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a protein disulfide isomerase-like polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a protein disulfide isomerase-like polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a protein disulfide isomerase-like polynucleotide can be used in a cell-based assay system. The protein disulfide isomerase-like polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a protein disulfide isomerase-like polypeptide, protein disulfide isomerase-like polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a protein disulfide isomerase-like polypeptide, or mimetics, agonists, antagonists, or inhibitors of a protein disulfide isomerase-like polypeptide activity. The compositions can be administered alone or

in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intraarterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethylcellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 150 mM histidine, 0.1%2% sucrose, and 27% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications

PDI has been associated with failure of potent anti-proliferative proteins such as TGFβ1 to inhibit proliferation of cancer cells. It is postulated that PDI acts by altering the binding of either the protein or its receptor. A therapeutic strategy targeted toward reducing the expression or activity of human protein disulfide isomerase-like enzyme, therefore, could enhance the anti-proliferative effective of anti-neoplastic compounds, particularly compounds which require specific protein conformations for maximum anti-proliferative effects for the treatment of cancer.

Cancer is a disease fundamentally caused by oncogenic cellular transformation. There are several hallmarks of transformed cells that distinguish them from their normal counterparts and underlie the pathophysiology of cancer. These include

uncontrolled cellular proliferation, unresponsiveness to normal death inducing signals (immortalization), increased cellular motility and invasiveness, increased ability to recruit blood supply through induction of new blood vessel formation (angiogenesis), genetic instability, and dysregulated gene expression. Various combinations of these aberrant physiologies, along with the acquisition of drug-resistance frequently lead to an intractable disease state in which organ failure and patient death ultimately ensue.

The advent of genomics-driven molecular target identification has opened up the possibility of identifying new cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for cancer patients. Thus, newly discovered tumor-associated genes and their products can be tested for their role(s) in disease and used as tools to discover and develop innovative therapies.

Genes playing important roles in any of the physiological processes outlined above can be characterized as cancer targets. Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized *in vitro* for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical activities. Agonists and/or antagonists of target protein activity can be identified in this manner and subsequently tested in cellular and *in vivo* disease models for anticancer activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

Human protein disulfide isomerase-like enzyme also can be regulated to treat hematological disorders and atherosclerosis. Atherosclerosis is associated with abnormally high level of PDI. Therefore, blocking the activity of human protein disulfide isomerase-like enzyme would be a viable approach to treating or preventing atherosclerosis.

Human protein disulfide isomerase-like enzyme also can be regulated to treat other cardiovascular disorders. Cardiovascular diseases include the following disorders of the heart and the vascular system: congestive heart failure, myocardial infarction, ischemic diseases of the heart, all kinds of atrial and ventricular arrhythmias, hypertensive vascular diseases and peripheral vascular diseases.

Heart failure is defined as a pathophysiologic state in which an abnormality of cardiac function is responsible for the failure of the heart to pump blood at a rate commensurate with the requirement of the metabolizing tissue. It includes all forms of pumping failure such as high-output and low-output, acute and chronic, right-sided or left-sided, systolic or diastolic, independent of the underlying cause.

Myocardial infarction (MI) is generally caused by an abrupt decrease in coronary blood flow that follows a thrombotic occlusion of a coronary artery previously narrowed by arteriosclerosis. MI prophylaxis (primary and secondary prevention) is included as well as the acute treatment of MI and the prevention of complications.

Ischemic diseases are conditions in which the coronary flow is restricted resulting in an perfusion which is inadequate to meet the myocardial requirement for oxygen. This group of diseases include stable angina, unstable angina and asymptomatic ischemia.

Arrhythmias include all forms of atrial and ventricular tachyarrhythmias (atrial tachycardia, atrial flutter, atrial fibrillation, atrio-ventricular reentrant tachycardia, preexcitation syndrome, ventricular tachycardia, ventricular flutter, ventricular fibrillation) as well as bradycardic forms of arrhythmias.

Hypertensive vascular diseases include primary as well as all kinds of secondary arterial hypertension (renal, endocrine, neurogenic, others). The genes may be used as drug targets for the treatment of hypertension as well as for the prevention of all

complications. Peripheral vascular diseases are defined as vascular diseases in which arterial and/or venous flow is reduced resulting in an imbalance between blood supply and tissue oxygen demand. It includes chronic peripheral arterial occlusive disease (PAOD), acute arterial thrombosis and embolism, inflammatory vascular disorders, Raynaud's phenomenon and venous disorders.

PDI is postulated to play a role in wound healing, because PDI is released from activated platelets. PDI may promote tissue remodeling and arrest of bleeding by reduction or isomerization of protein bonds. By enhancing the expression of human protein disulfide isomerase-like enzyme, it may be possible to increase the levels of properly folded proteins that are important in facilitating would healing.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a human protein disulfide isomerase-like polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above described screening assays for treatments as described herein.

A reagent which affects human protein disulfide isomerase-like enzyme activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce human protein disulfide isomerase-like enzyme activity. The reagent preferably binds to an expression product of a human protein disulfide isomerase-like enzyme gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which

have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988); Wu *et al.*, *J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem. 266*, 338-42 (1991).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases human protein disulfide isomerase-like enzyme activity relative to the human protein disulfide isomerase-like enzyme activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a human protein disulfide isomerase-like enzyme gene or the activity of a human protein disulfide isomerase-like polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a human protein disulfide isomerase-like enzyme gene or the activity of a human protein disulfide isomerase-like polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to human protein disulfide isomerase-like enzyme-specific mRNA, quantitative RT-PCR, immunologic detection of a human protein disulfide

isomerase-like polypeptide, or measurement of human protein disulfide isomerase-like enzyme activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic Methods

Human protein disulfide isomerase-like enzyme also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences which encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding protein disulfide isomerase-like enzyme in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing

primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al., Science 230*, 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al., Proc. Natl. Acad. Sci. USA 85*, 43974401, 1985). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of a protein disulfide isomerase-like enzyme also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following

specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of protein disulfide isomerase-like enzyme activity

The polynucleotide of SEQ ID NO:1 or SEQ ID NO:28 is inserted into the expression vector pCEV4 and the expression vector pCEV4-protein disulfide isomerase-like enzyme polypeptide obtained is transfected into human embryonic kidney 293 cells. From these cells extracts are obtained and the protein disulfide isomerase-like enzyme activity is measured by it's ability to refold scrambled Rnase A, which in turn will regain its ability to cleave highly polymerized yeast RNA (Lambert and Freedman, 1983). The cell extract (2,0µg/ml) is dissolved in 100 mM sodium phosphate, pH 7,5, 2 mM glutathione, and 0,2 mM reduced glutathione. The reaction is started by the addition of 200 µg of scrambled Rnase A (Hillson et al., 1984). Protein disulfide isomerase-like enzyme activity is determined in the presence of recombinant calretioulin, purified N-, P-, and C-domains of calreticulin, and in the presence of cardiac calsequestrin. For the assay, the proteins are used in 1:1 molar ratio. The A260/A280 ratio over a period of 60 min is measured as indicative of the extent of cleavage of yeast RNA by refolded Rnase A. It is shown that the polypeptide of SEQ ID NO:2 and SEQ ID NO:29 has a protein disulfide isomerase-like enzyme activity.

EXAMPLE 2

Expression of recombinant human protein disulfide isomerase-like enzyme

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human protein disulfide isomerase-like polypeptides in yeast. The protein disulfide isomerase-like enzyme-encoding DNA

sequence is derived from SEQ ID NO:1 or SEQ ID NO:28. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5' end an initiation codon and at its 3' end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/mdHis6 vector is used to transform the yeast

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (NiNTAResin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human protein disulfide isomerase-like polypeptide is obtained.

EXAMPLE 3

Identification of a test compound which decreases protein disulfide isomerase-like activity

Human protein disulfide isomerase-like enzyme activity is measured by renaturation of inactivated RNase. Briefly, RNase is inactivated by formation of randomly mismatched or scrambled disulfide bonds. Deactivated RNase renatures slowly in the presence of a thiol reagent. Human protein disulfide isomerase-like enzyme accelerates the renaturation process. The amount of renatured RNase is measured by its ability to digest RNA. RNA digestion is measured by an increase in absorbance at 260 nm.

A test compound which decreases protein disulfide isomerase-like activity relative to control levels by at least 20% is identified as a protein disulfide isomerase-like enzyme inhibitor.

EXAMPLE 4

Identification of test compounds that bind to protein disulfide isomerase-like polypeptides

Purified protein disulfide isomerase-like polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Protein disulfide isomerase-like polypeptides comprise the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a protein disulfide isomerase-like polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a protein disulfide isomerase-like polypeptide.

EXAMPLE 5

Identification of a test compound which decreases protein disulfide isomerase-like enzyme gene expression

A test compound is administered to a culture of human cells transfected with a protein disulfide isomerase-like enzyme expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells which have not been transfected is incubated for the same time without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 Fg total RNA and hybridized with a ^{32}P-labeled protein disulfide isomerase-like enzyme-specific probe at 65° C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO:1 or SEQ ID NO:28. A test compound which decreases the protein disulfide isomerase-like enzyme-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of protein disulfide isomerase-like enzyme gene expression.

EXAMPLE 6

Proliferation inhibition assay: Antisense oligonucleotides suppress the growth of cancer cell lines in the presence of TGFβ1

A human pancreatic carcinoma line which does not downregulate the expression of human protein disulfide isomerase-like enzyme in response to TGFβ1 and which is resistant to the anti-proliferative effects of TGFβ1 is used for testing. Cells are cultured in RPMI-1640 with 10-15% fetal calf serum at a concentration of 10,000 cells per milliliter in a volume of 0.5 ml and kept at 37 °C in a 95% air/5%CO_2 atmosphere.

Phosphorothioate oligoribonucleotides are synthesized on an Applied Biosystems Model 380B DNA synthesizer using phosphoroamidite chemistry. A sequence of 24 bases complementary to the nucleotides at position 1 to 24 of SEQ ID NO:1 or SEQ ID NO:28 is used as the test oligonucleotide. As a control, another (random)

sequence is used: 5'-TCA ACT GAC TAG ATG TAC ATG GAC-3'. Following assembly and deprotection, oligonucleotides are ethanol-precipitated twice, dried, and suspended in phosphate buffered saline at the desired concentration. Purity of the oligonucleotides is tested by capillary gel electrophoresis and ion exchange HPLC. The purified oligonucleotides are added to the culture medium at a concentration of 10 μM once per day for seven days.

The addition of the test oligonucleotide for seven days together with TGFβ1 results in significantly reduced expression of human protein disulfide isomerase-like enzyme as determined by Western blotting. This effect is not observed with the control oligonucleotide. After 3 to 7 days, the number of cells in the cultures is counted using an automatic cell counter. The number of cells in cultures treated with the test oligonucleotide (expressed as 100%) is compared with the number of cells in cultures treated with the control oligonucleotide. The number of cells in cultures treated with the test oligonucleotide is not more than 30% of control, indicating that the inhibition of human protein disulfide isomerase-like enzyme has an anti-proliferative effect on cancer cells.

EXAMPLE 7

Quantitative expression profiling of human protein disulfide isomerase-like enzyme

Expression profiling is based on a quantitative polymerase chain reaction (PCR) analysis, also called kinetic analysis, first described in Higuchi et al., 1992 and Higuchi et al., 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies. Using this technique, the expression levels of particular genes, which are transcribed from the chromosomes as messenger RNA (mRNA), are measured by first making a DNA copy (cDNA) of the mRNA, and then performing quantitative PCR on the cDNA, a method called quantitative reverse transcription-polymerase chain reaction (quantitative RT-PCR).

Quantitative RT-PCR analysis of RNA from different human tissues was performed to investigate the tissue distribution of human protein disulfide isomerase-like enzyme, Sphingosine kinase-like mRNA. In most cases, 25 .µg of total RNA from various tissues (including Human Total RNA Panel I-V, Clontech Laboratories, Palo Alto, CA, USA) was used as a template to synthsize first-strand cDNA using the SUPERSCRIPT™ First-Strand Synthesis System for RT-PCR (Life Technologies, Rockville , MD, USA). First-strand cDNA synthesis was carried out according to the manufacturer's protocol using oligo (dT) to hybridize to the 3' poly A tails of mRNA and prime the synthesis reaction. Approximately 10 ng of the first-strand cDNA was then used as template in a polymerase chain reaction. In other cases, 10 ng of commercially available cDNAs (Human Immune System MTC Panel and Human Blood Fractions MTC Panel, Clontech Laboratories, Palo Alto, CA, USA) were used as template in a polymerase chain reaction. The polymerase chain reaction was performed in a LightCycler (Roche Molecular Biochemicals, Indianapolis, IN, USA), in the presence of the DNA-binding fluorescent dye SYBR Green I which binds to the minor groove of the DNA double helix, produced only when double-stranded DNA is successfully synthesized in the reaction. Upon binding to double-stranded DNA, SYBR Green I emits light that can be quantitatively measured by the LightCycler machine. The polymerase chain reaction was carried out using oligonucleotide primers LBRI059-L2:

GGACACATTCACTGTGGCAAGG
and LBRI059-R2:
TGGAAGCAGGG TCTGAGTGATG

and measurements of the intensity of emitted light were taken following each cycle of the reaction when the reaction had reached a temperature of 81 degrees C. Intensities of emitted light were converted into copy numbers of the gene transcript per nanogram of template cDNA by comparison with simultaneously reacted standards of known concentration.

To correct for differences in mRNA transcription levels per cell in the various tissue types, a normalization procedure was performed using similarly calculated expression levels in the various tissues of five different housekeeping genes: glyceraldehyde-3-phosphatase (G3PDH), hypoxanthine guanine phophoribosyl transferase (HPRT), beta-actin, porphobilinogen deaminase (PBGD), and beta-2-microglobulin. The level of housekeeping gene expression is considered to be relatively constant for all tissues and therefore can be used as a gauge to approximate relative numbers of cells per µg of total RNA used in the cDNA synthesis step. Except for the use of a slightly different set of housekeeping genes and the use of the LightCycler system to measure expression levels, the normalization procedure was similar to that described in the RNA Master Blot User Manual, Apendix C (1997, Clontech Laboratories, Palo Alto, CA, USA). In brief, expression levels of the five housekeeping genes in all tissue samples were measured in three independent reactions per gene using the LightCycler and a constant amount (25 µg) of starting RNA. The calculated copy numbers for each gene, derived from comparison with simultaneously reacted standards of known concentrations, were recorded and the mean number of copies of each gene in all tissue samples was determined. Then for each tissue sample, the expression of each housekeeping gene relative to the mean was calculated, and the average of these values over the five housekeeping genes was found. A normalization factor for each tissue was then calculated by dividing the final value for one of the tissues arbitrarily selected as a standard by the corresponding value for each of the tissues. To normalize an experimentally obtained value for the expression of a particular gene in a tissue sample, the obtained value was multiplied by the normalization factor for the tissue tested. This normalization method was used for all tissues except those derived from the Human Blood Fractions MTC Panel, which showed dramatic variation in some housekeeping genes depending on whether the tissue had been activated or not. In these tissues, normalization was carried out with a single housekeeping gene, beta-2-microglobulin.

Results are given in Figures 35 and 36, showing the experimentally obtained copy numbers of mRNA per 10 ng of first-strand cDNA on the left and the normalized values on the right. RNAs used for the cDNA synthesis, along with their supplier and catalog numbers are shown in tables 1 and 2.

Human protein disulfide isomerase-like enzyme is a widely distributed, highly expressed molecule, indicating that its function may be fundamental to common activities of a wide variety of cell types. As a member of the protein disulfide isomerase family, this function would most likely be in helping different proteins assume their proper three-dimensional configuration.

Although the expression of human protein disulfide isomerase-like enzyme is high in most tissues, its expression in the lung is 1.5 to 2 times that seen in any other tissue. This suggests that the lung may produce a comparatively large amount of specific protein types, for example proteins lining or secreted into the lung airways, that require human protein disulfide isomerase-like enzyme activity to allow them to fold or combine with other molecules before assuming an active conformation. Regulating the activity of human protein disulfide isomerase-like enzyme in cells or tissues such as lung that produce human protein disulfide isomerase-like enzyme - activity-dependent proteins, therefore, may be beneficial in pathogenic states where overproduction or underproduction of active protein leads to disease.

Table 1: Material used to determine the expression in whole body tissues.

Tissue	Supplier	Panel name and catalog number
1. brain	Clontech	Human Total RNA Panel I, K4000-1
2. heart	Clontech	Human Total RNA Panel I, K4000-1
3. kidney	Clontech	Human Total RNA Panel I, K4000-1
4. liver	Clontech	Human Total RNA Panel I, K4000-1
5. lung	Clontech	Human Total RNA Panel I, K4000-1
6. trachea	Clontech	Human Total RNA Panel I, K4000-1
7. bone marrow	Clontech	Human Total RNA Panel II, K4001-1
8. colon	Clontech	Human Total RNA Panel II, K4001-1
9. small intestine	Clontech	Human Total RNA Panel II, K4001-1
10. spleen	Clontech	Human Total RNA Panel II, K4001-1
11. stomach	Clontech	Human Total RNA Panel II, K4001-1
12. thymus	Clontech	Human Total RNA Panel II, K4001-1
13. mammary gland	Clontech	Human Total RNA Panel III, K4002-1
14. skeletal muscle	Clontech	Human Total RNA Panel III, K4002-1
15. prostate	Clontech	Human Total RNA Panel III, K4002-1
16. testis	Clontech	Human Total RNA Panel III, K4002-1
17. uterus	Clontech	Human Total RNA Panel III, K4002-1
18. cerebellum	Clontech	Human Total RNA Panel IV, K4003-1
19. fetal brain	Clontech	Human Total RNA Panel IV, K4003-1
20. fetal liver	Clontech	Human Total RNA Panel IV, K4003-1
21. spinal cord	Clontech	Human Total RNA Panel IV, K4003-1
22. placenta	Clontech	Human Total RNA Panel IV, K4003-1
23. adrenal gland	Clontech	Human Total RNA Panel V, K4004-1
24. pancreas	Clontech	Human Total RNA Panel V, K4004-1
25. salivary gland	Clontech	Human Total RNA Panel V, K4004-1
26. thyroid	Clontech	Human Total RNA Panel V, K4004-1

Table 2: Material used to determine the expression in the blood and lung tissues.

Tissue	Supplier	Panel name and catalog number
1. lymph node	Clontech	Human Immune System MTC Panel, K1426-1
2. peripheral blood leukocytes	Clontech	Human Immune System MTC Panel, K1426-1
3. tonsil	Clontech	Human Immune System MTC Panel, K1426-1
4. peripheral blood mononuclear cells	Clontech	Human Blood Fractions MTC Panel, K1428-1
5. peripheral blood mononuclear cells – activated	Clontech	Human Blood Fractions MTC Panel, K1428-1
6. T-cell (CD8+)	Clontech	Human Blood Fractions MTC Panel, K1428-1
7. T-cell (CD8+) - activated	Clontech	Human Blood Fractions MTC Panel, K1428-1
8. T-cell (CD4+)	Clontech	Human Blood Fractions MTC Panel, K1428-1
9. T-cell (CD4+) – activated	Clontech	Human Blood Fractions MTC Panel, K1428-1
10. B-cell (CD19+)	Clontech	Human Blood Fractions MTC Panel, K1428-1
11. B-cell (CD19+) – activated	Clontech	Human Blood Fractions MTC Panel, K1428-1
12. Monocytes (CD14+)	Clontech	Human Blood Fractions MTC Panel, K1428-1
13. Th1 clone	In-house	
14. Th2 clone	In-house	
15. neutrophil	In-house	
16. neutrophil	In-house	
17. Normal Bronchial/Tracheal Epithelial Cells	In-house	
18. Normal Bronchial/Tracheal smooth muscle cell	In-house	
19. Normal lung fibroblast	In-house	
20. Microvascular Endothelial cell	In-house	
21. U937	In-house	
22. RAMOS	In-house	
23. Jurkat	In-house	
24. HelaS3	In-house	
25. IMR-90	In-house	
26. HEK293	In-house	

CLAIMS

1. An isolated polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide and being selected from the group consisting of:

   a) a polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide comprising an amino acid sequence selected form the group consisting of:

      amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2,
      amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29,
      the amino acid sequence shown in SEQ ID NO:2; and
      the amino acid sequence shown in SEQ ID NO:29.

   b) a polynucleotide comprising the sequence of SEQ ID NO:1 or SEQ ID NO:28;

   c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

   d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

   e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a to (d).

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified protein disulfide isomerase-like enzyme polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a protein disulfide isomerase-like enzyme polypeptide, wherein the method comprises the following steps:

   a) culturing the host cell of claim 3 under conditions suitable for the expression of the protein disulfide isomerase-like enzyme polypeptide; and
   b) recovering the protein disulfide isomerase-like enzyme polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a protein disulfide isomerase-like enzyme polypeptide in a biological sample comprising the following steps:

   a) hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

   b) detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a protein disulfide isomerase-like enzyme polypeptide of claim 4 comprising the steps of:

contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the protein disulfide isomerase-like enzyme polypeptide.

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a test compound with any protein disulfide isomerase-like enzyme polypeptide encoded by any polynucleotide of claim1;

detecting binding of the test compound to the protein disulfide isomerase-like enzyme polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a protein disulfide isomerase-like enzyme.

11. A method of screening for agents which regulate the activity of a protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a test compound with a protein disulfide isomerase-like enzyme polypeptide encoded by any polynucleotide of claim 1; and

detecting a protein disulfide isomerase-like enzyme activity of the polypeptide, wherein a test compound which increases the protein disulfide isomerase-like enzyme activity is identified as a potential therapeutic agent for increasing the activity of the protein disulfide isomerase-like enzyme, and wherein a test compound which decreases the protein disulfide isomerase-like enzyme activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the protein disulfide isomerase-like enzyme.

12. A method of screening for agents which decrease the activity of a protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of protein disulfide isomerase-like enzyme.

13. A method of reducing the activity of protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any protein disulfide isomerase-like enzyme polypeptide of claim 4, whereby the activity of protein disulfide isomerase-like enzyme is reduced.

14. A reagent that modulates the activity of a protein disulfide isomerase-like enzyme polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15. A pharmaceutical composition, comprising:

the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the pharmaceutical composition of claim 15 for modulating the activity of a protein disulfide isomerase-like enzyme in a disease.

17. Use of claim 16 wherein the disease is cancer, diabetes and obesity.

18. A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29.

19. The cDNA of claim 18 which comprises SEQ ID NO:1 or SEQ ID NO:28.

20. The cDNA of claim 18 which consists of SEQ ID NO:1 or SEQ ID NO:28.

21. An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29.

22. The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO:1 or SEQ ID NO:28.

23. A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29.

24. The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO:1 or SEQ ID NO:28.

25. A purified polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29.

26. The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29.

27. A fusion protein comprising a polypeptide having the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29.

28. A method of producing a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29, comprising the steps of:

culturing a host cell comprising an expression vector which encodes the polypeptide under conditions whereby the polypeptide is expressed; and

isolating the polypeptide.

29. The method of claim 28 wherein the expression vector comprises SEQ ID NO:1 or SEQ ID NO:28.

30. A method of detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29, comprising the steps of:

hybridizing a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1 or SEQ ID NO:28 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and

detecting the hybridization complex.

31. The method of claim 30 further comprising the step of amplifying the nucleic acid material before the step of hybridizing.

32. A kit for detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29, comprising:

a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1 or SEQ ID NO:28; and

instructions for the method of claim 30.

33. A method of detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29, comprising the steps of:

contacting a biological sample with a reagent that specifically binds to the polypeptide to form a reagent-polypeptide complex; and

detecting the reagent-polypeptide complex.

34. The method of claim 33 wherein the reagent is an antibody.

35. A kit for detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29, comprising:

an antibody which specifically binds to the polypeptide; and

instructions for the method of claim 33.

36. A method of screening for agents which can modulate the activity of a human protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2 or wich are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29 and (2) the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29; and

detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential agent for regulating activity of the human protein disulfide isomerase-like enzyme.

37. The method of claim 36 wherein the step of contacting is in a cell.

38. The method of claim 36 wherein the cell is *in vitro*.

39. The method of claim 36 wherein the step of contacting is in a cell-free system.

40. The method of claim 36 wherein the polypeptide comprises a detectable label.

41. The method of claim 36 wherein the test compound comprises a detectable label.

42. The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.

43. The method of claim 36 wherein the polypeptide is bound to a solid support.

44. The method of claim 36 wherein the test compound is bound to a solid support.

45. A method of screening for agents which modulate an activity of a human protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:2 or which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO:29 and (2) the amino acid sequence shown in SEQ ID NO:2 or in SEQ ID NO:29; and

detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human protein disulfide isomerase-like enzyme, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential agent for decreasing the activity of the human protein disulfide isomerase-like enzyme.

46. The method of claim 45 wherein the step of contacting is in a cell.

47. The method of claim 45 wherein the cell is *in vitro*.

48. The method of claim 45 wherein the step of contacting is in a cell-free system.

49. A method of screening for agents which modulate an activity of a human protein disulfide isomerase-like enzyme, comprising the steps of:

contacting a test compound with a product encoded by a polynucleotide which comprises the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:28; and

detecting binding of the test compound to the product, wherein a test compound which binds to the product is identified as a potential agent for regulating the activity of the human protein disulfide isomerase-like enzyme.

50. The method of claim 49 wherein the product is a polypeptide.

51. The method of claim 49 wherein the product is RNA.

52. A method of reducing activity of a human protein disulfide isomerase-like enzyme, comprising the step of:

contacting a cell with a reagent which specifically binds to a product encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:28, whereby the activity of a human protein disulfide isomerase-like enzyme is reduced.

53. The method of claim 52 wherein the product is a polypeptide.

54. The method of claim 53 wherein the reagent is an antibody.

55. The method of claim 52 wherein the product is RNA.

56. The method of claim 55 wherein the reagent is an antisense oligonucleotide.

57. The method of claim 56 wherein the reagent is a ribozyme.

58. The method of claim 52 wherein the cell is *in vitro*.

59. The method of claim 52 wherein the cell is *in vivo*.

60. A pharmaceutical composition, comprising:

a reagent which specifically binds to a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29; and

a pharmaceutically acceptable carrier.

61. The pharmaceutical composition of claim 60 wherein the reagent is an antibody.

62. A pharmaceutical composition, comprising:

a reagent which specifically binds to a product of a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:28; and a pharmaceutically acceptable carrier.

63. The pharmaceutical composition of claim 62 wherein the reagent is a ribozyme.

64. The pharmaceutical composition of claim 62 wherein the reagent is an antisense oligonucleotide.

65. The pharmaceutical composition of claim 62 wherein the reagent is an antibody.

66. A pharmaceutical composition, comprising:

an expression vector encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:29; and

a pharmaceutically acceptable carrier.

67. The pharmaceutical composition of claim 66 wherein the expression vector comprises SEQ ID NO:1 or SEQ ID NO:28.

68. A method of treating a protein disulfide isomerase-like enzyme disfunction related disease, wherein the disease is selected from cancer, diabetes and obesity, comprising the step of:

administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human protein disulfide isomerase-like enzyme, whereby symptoms of the protein disulfide isomerase-like enzyme disfunction related disease are ameliorated.

69. The method of claim 68 wherein the reagent is identified by the method of claim 36.

70. The method of claim 68 wherein the reagent is identified by the method of claim 45.

71. The method of claim 68 wherein the reagent is identified by the method of claim 49.

Fig. 1

```
atgaagctca ccctagagtc
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct
tcctactgtc ttgttttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga
cgtccctatc acccttccaa acctgttgag gttcattttg catcactcag accctgcttc
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg
caagctgcag gagctggccg atgcctcaga aaacctcctt accgagaaca cgtggctcaa
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg
cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa
caggacagac taa
```

Fig. 2

MKLTLESFIQ NFSVLYSPLK RHLIGSGSAQ FPSQHLITEV TTDTFWEVVL QKQDVLLLYY
APWCGFCPSL NHIFIQLARN LPMDTFTVAR IDVSQNDLPW EFMVDRLPTV LFFPCNRKDL
SVKYPEDVPI TLPNLLRFIL HHSDPASSPQ NVANSPTKEC LQSEAVLQRG HISHLEREIQ
KLRAEISSLQ RAQVQVESQL SSARRDEHRL RQQQRALEEQ HSLLHAHSEQ LQALYEQKTR
ELQELARKLQ ELADASENLL TENTWLKILV ATMERKLEGR DGAESLAAQR EVHPKQPEPS
ATPQLPGSSP PPANVSATLV SERNKENRTD

Fig. 3

MKISNLLAAF LAFSGGFFCA SAEVPKVNKE GLNELITADK VLMVKFYAPW CGHCKALAPE
YESAADELEK DGISLVEVDC TEEGDLCSEY SIRGYPTLNV FKNGKQISQY SGPRKHDALV
KYMRKQLLPT VKPISKDTLE NFVEKADDLA VVAFFKDQKL NDTYTEVAEV MKDDFVFAAS
DDKELAKSLG SNFPGIVAFT KDAAQDSDKL VYTGDWDPAS IADFIGVSSI PLLDELNQMT
FGKYQQSGLP LGIIFYNSTE SRDELYDVFQ PLAKKYQDTL RFAFLDAVRY GAVAKQMNVE
SDWPAFVIAN LKSMLKYPFP TTELTAKAMT KFVGDFVDGK LQPKIKSQPI PESQEDLVVL
VADNFDDIVM DETKDVLVEF YAPWCGHCKN LAPTYEKLAE EYSDDSNVVV AKIDATENDI
SVSISGFPTI MFFKANDKVN PVRYEGDRTL EDLSAFIDKH ASFEPIKKEK ESVPAPDLED
QVAVEDEMAD EL

Fig. 4

ssvvvvltdenFdeevlkaksdkpVLVdFyApWCGpC

Fig. 5

```
ttttaacttc ctatagcccc ttcagctact acactgcatg ttgcaggacc ataagcaggg
gtgtgtcagg cttcatcgac tctgaacaag gtgtctttga agcccctact gttgcatttt
cttcccttga gaagaaatgt gaggttgatg ccccaagctc cgttcctcac attgaggaga
acaggtatct ctttccagaa gtggacatga ctagcacaaa cttcacaggc ctgagctgca
gaaccaacaa gactctcaac atctaccttt tggattcaaa tttgtttggt tatatgcaga
gagactgggt gctccgagct ccactcaggt gaaagaattt gcggcaattg ttgacgtgaa
agaagaatct cattacatct tggatccaaa gcaagcactg atgaagctca ccctagagtc
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct
tcctactgtc ttgtttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga
cgtccctatc acccttccaa acctgttgag gttcatttg catcactcag accctgcttc
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg
caagctgcag gagctggccg atgcctcaga aaacctcctt accgagaaca cgtggctcaa
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg
cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa
caggacagac taacttttta aatgatatga agaaatcaga ggtgaaaatt gtacattggg
aatatattta tgcaaatttt attgaaattt attgtaaata aagattttct cagtggtcta
gaaaatcaaa aaaaaaaaaa aaaaaaaaaa aa
```

Fig. 6

```
ttttaacttc ctatagcccc ttcagctact acactgcatg ttgcaggacc ataagcaggg
gtgtgtcagg cttcatcgac tctgaacaag gtgtctttga agcccctact gttgcatttt
cttcccttga gaagaaatgt gaggttgatg ccccaagctc cgttcctcac attgaggaga
acaggtatct ctttccagaa gtggacatga ctagcacaaa cttcacaggc ctgagctgca
gaaccaacaa gactctcaac atctaccttt tggattcaaa tttgtttggt tatatgcaga
gagactgggt gctccgagct ccactcaggt gaaagaattt gcggcaattg ttgacgtgaa
agaagaatct cattacatct tggatccaaa gcaagcactg atgaagctca ccctagagtc
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct
tcctactgtc ttgtttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga
cgtccctatc acccttccaa acctgttgag gttcattttg catcactcag accctgcttc
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg
caagctgcag gagctggccg atgcctcaga aaacctcctt accgagaaca cgtggctcaa
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg
```

Fig. 6 (continued)

```
cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa
caggacagac taacttttta aatgatatga agaaatcaga ggtgaaaatt gtacattggg
aatatattta tgcaaatttt attgaaattt attgtaaata aagattttct cagtggtcta
gaaaatcaaa aaaaaaaaaa aaaaaaaaaa aa
```

Fig. 7

```
ttttaacttc ctatagcccc ttcagctact acactgcatg ttgcaggacc ataagcaggg
gtgtgtcagg cttcatcgac tctgaacaag gtgtctttga agcccctact gttgcatttt
cttcccttga gaagaaatgt gaggttgatg ccccaagctc cgttcctcac attgaggaga
acaggtatct ctttccagaa gtggacatga ctagcacaaa cttcacaggc ctgagctgca
gaaccaacaa gactctcaac atctaccttt tggattcaaa tttgtttggt tatatgcaga
gagactgggt gctccgagct ccactcaggt gaaagaattt gcggcaattg ttgacgtgaa
agaagaatct cattacatct tggatccaaa gcaagcactg atgaagctca ccctagagtc
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct
tcctactgtc ttgttttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga
cgtccctatc acccttccaa acctgttgag gttcattttg catcactcag accctgcttc
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg
caagctgcag gagctggccg atgcctcaga aaacctcctt accgagaaca cgtggctcaa
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg
```

Fig. 7 (continued)

cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa
caggacagac taacttttta aatgatatga agaaatcaga ggtgaaaatt gtacattggg
aatatattta tgcaaatttt attgaaattt attgtaaata aagattttct cagtggtcta
gaaaatcaaa aaaaaaaaaa aaaaaaaaaa aa

Fig. 8

ACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATATTCCC
AATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCTGTTC
TCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCTGCCA
GGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGCCGCC
AGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGATCTTG
AGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTTGCGG
GCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACTGTGT
GCGTGGAGCAGGCTGTGCTGCTCTTCCAGGGCCCGCTGCTGCTGCCGCAGCCGGTGCTCA
TCTCTGCGGGCACTGGAGAGCTGGGACTCCACCTGCACTTGTGCTCGCTGGAGGCTGCTT
ATTTCTGCTCTCAGTTTCTGGATCTCTCTCTCCAAGTGGGAGATGTGCCCCCGCTGTAAG
ACTGCCTCGCTCTGAAGACACTCCTTGGTAGGAGAGTTAGCCACATTCTGGGGGCTGGAA
AGCAGGGTCTGAGTGATGCAAAATGAACCTCAACAGGTTTGGAAGGGTGATGGGGACGTC
TTCGGGGTATTTCACACTTAGGGTCCTTTCTGTTGCAGG

Fig. 9

TAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATATT
CCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCTG
TTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCTG
CCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGCC
GCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGATC
TTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTTG
CGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACTG
TGTGCGTGGAGCAAGCTGTGCTGCTCTTCCAGGGCCCGCTGCTGCTGCCGCAGCCGGTGC
TCATCTCTGCGGGCACTGGAGAGCTGGGACTCCACCTGCACTTGTGCTCGCTGGAGGCTG
CTTATTTCTGCTCTCAGTTTCTGGATCTCTCTCTCCAAGTGGGAGATGTGCCCCCGCTGT
AAGACTGCCTCGCTCTGAAGACACTCCTTGGTAGGAGAGTTAGCCACATTCTNGGGGCTG
GAAGCAGGGTCTGAATGATGCAAAATGAACCCTAACAGGTTTGGAAAGGTGATGGGGAGG
TCTTCCGGGGTATTTCACACTTAGGTCCCTTCTGTTGGACGGAAAAAAAACAGCG

Fig. 10

TAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATATT
CCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCTG
TTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCTG
CCAGGGAGCTGGGGGGTGGCTGAGGGCTCANGCTGCTTGGGGTGGACCTCTCTCTGGGCC
GCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGATC
TTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTTG
CGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACTG
TGTGCGTGGAGCAGGCTGTGCTGCTCTTCCAGGGCCCGCTGCTGCTGCCGCAGCCGGTGC
TCATCTCTGCGGGCACTGGAGAGCTGGGACTCCACCTGCACTTGTGCTCGCTG

Fig. 11

CGAGCTCCACTCAGGTGAAAGAATTTGCGGCAATTGTTGACGTGAAAGAAGAATCTCATT
ACATCTTGGATCCAAAGCAAGCACTGATGAAGCTCACCCTAGAGTCTTTTATTCAAAACT
TCAGCGTTCTCTATAGTCCCTTGAAAAGGCATCTCATTGGAAGTGGCTCTGCCCAGTTCC
CGTCTCAGCATTTAATCACTGAAGTGACAACTGATACCTTTTGGGAAGTAGTCCTTCAAA
AACAGGACGTTCTCCTGCTCTATTACGCTCCGTGGTGCGGCTTCTGTCCATCCCTCAATC
ACATCTTCATCCAGCTAGCTCGGAACCTGCCCATGGACACATTCACTGTGGCAAGGATTG
ACGTGTCTCANAATGACCTTCCTTGGGAATTTATGGTCGATCGTCTNCCTACTGTCTTGT
TTTTTCCCTGCAACAGAAAGGACCTAAGTGTGAAATACCCCGAAGACGTCCTCATCACCC
TTCCAAACCTGGTGAGGTTCATTTTGCATCACTCAGACCCTGCTTCCAGCCCCCAGATGT
GGC

Fig. 12

TAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATATT
CCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCTG
TTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCTG
CCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGCC
GCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGATC
TTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTTG
CGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACTG
TGTGCGTGGAGCAGGCTGTGCTGCTCTTCCAGGGCCCGCTGCTGCTGCCGCAGCCGGTGC
TCATCTCTGCG

Fig. 13

CTGAAGTGACAACTGATACCTTTTGGGAAGTAGTCCTTCAAAAACAGGACGTTCTCCTGC
TCTATTACGCTCCGTGGTGCGGCTTCTGTCCATCCCTCAATCACATCTTCATCCAGCTAG
CTCGGAACCTGCCCATGGACACATTCACTGTGGCAAGGATTGACGTGTCTCAGAATGACC
TTCCTTGGGAATTTATGGTCGATCGTCTTCCTACTGTCTTGTTTTTCCCTGCAACAGAA
AGGACCTAAGTGTGAAATACCCCGAAGACGTCCCCATCACCCTTCCAAACCTGTTGAGGT
TCATTTTGCATCACTCAGACCCTGCTTCCAGCCCCCAGAATGTGGCTAACTCTCCTACCA
AGGAGTGTCTTCAGAGCGAGGCAGTCTTACAGCGGGGCACATNTNCCACTTGGAGAGAG
AGATCCAGAAACTGAGAGCAGAAATAAGCAGCCTNCAGCGAGCACAAGTGCAGGTGGAGT
CCCAGCTC

Fig. 14

CTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATAT
TCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCT
GTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCT
GCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGC
CGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGAT
CTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTT
GCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACT
GTGTGCGTGGAGCAGGCTGTGCTGCTCTTCCAGGGCCCGCTGCTGCTG

Fig. 15

TTTTTTNATGTTTTTTTTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAA
ATTTGCATAAATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAA
AAAGTTAGTCTGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGG
CAGGTGGAGGGGAGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGT
GGACCTCTCTCTGGGCCGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCA
TGGTCGCCACCAGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGG
CCAGCTCCTGCAGCTTGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGG
CCTGCAGCTGCTCACTGTGTGCGTGGAGCAGGCTGTGCT

Fig. 16

TCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATA
TTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCC
TGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGC
TGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGG
CCGCCACGGTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGA
TCTTGAGCCACGTGTTCTCCGTAAGAGNNTTTTCTGAGGCATCGGCCAGCTCCTGCAGCT
TGCGAGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCAC
TGTGTGCGTGGAGCAGGCTGTGCTGCTCTTCCAGGGCCCGCTGCTGCTGCCG

Fig. 17

TCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATA
TTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCC
TGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGC
TGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGG
CCGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGA
TCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCT
TGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCAC
TGTGTGCGTGGAGC

Fig. 18

GATTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAA
ATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTC
TGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGG
GGAGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCANGCTGCTTGNGGTGGACCTNTCT
CTGGGCCGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCAC
CAGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTG
CAGCTTGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTG
CTCACTGTGTGCGTGGAGCA

Fig. 19

AGACCATTTAGTAAATCTTTATTTACAATAAATTTCAATGAAATTTGCATAAATATATTC
CCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCTGT
TCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCTGC
CAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGCCG
CCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGATCT
TGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTTGC
GGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACTGT
GTGCGTGGAGCAGGCTGTGCTGCTCT

Fig. 20

TTTTTTTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCA
TAAATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTA
GTCTGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGG
AGGGGAGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTC
TCTCTGGGCCGCCACGTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCC
ACCAGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCC
TGCAGCTTGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGC
TGCTCACTGTGTGCGTGGAG

Fig. 21

TTTTTTTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCA
TAAATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTA
GTCTGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGG
AGGGGAGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTC
TCTCTGGGCCGCCACGTTTACAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGC
CACCAGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTC
CTGCAGCTTGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAG
CTGCTCACTGTGTGCGTGGAG

Fig. 22

CTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATAT
TCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCT
GTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCT
GCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGC
CGCCACGTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGATC
TTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTTG
CGTGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGCTCACTG
TGTGCGTGGAGCAGGCT

Fig. 23

TTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAAT
ATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTG
TCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGG
AGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCT
GGGCCGCCACGTTATCAGCCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACC
AGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGC
AGCTTGCGAGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGCCTGCAGCTGC
TCACTGTGTGCGTGGAG

Fig. 24

TTTTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAA
ATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTC
TGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGACG
GGAGCAGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCT
CTGGGCCGCCACGTTTCTAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCAC
CAGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTG
CAGCTTGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATAGAGGGGCTGAAGCTG
CTCACTGTGTGCGTGGAGAAG

Fig. 25

TTTTTTTTTTTTTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAA
TTTGCATAAATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAA
AAGTTAGTCTGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGC
AGGTGGAGGGGAGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTG
GACCTCTCTCTGGGCCGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCAT
GGTCGCCACCAGGATCTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGC
CAGCTCCTGCAGCTTGCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATA

Fig. 26

TTTTTTTTTCTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCA
TAAATATATTCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTA
GTCTGTCCTGTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGG
AGGGGAGCTGCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTC
TCTCTGGGCCGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGC
CACCANGATCTTGAGCCACGTGTTCTCGGTAGGGAGGTTTTCTGAGGCATCGGCCAGCTC
CTGCAGCTTGCGGGCCAGCTCCTGCACTTCAACGTGTCTTCTGCTCATTAGAGGGCCTGC
AGCTGCTCACTGTGTTGCGTGGAGCAAGGCCTGTGCTTG

Fig. 27

CTAGACCACTGAGAAAATCTTTATTTACAATAAATTTCAATAAAATTTGCATAAATATAT
TCCCAATGTACAATTTTCACCTCTGATTTCTTCATATCATTTAAAAAGTTAGTCTGTCCT
GTTCTCCTTATTCCTTTCAGACACCAGTGTGGCGCTGACATTGGCAGGTGGAGGGGAGCT
GCCAGGGAGCTGGGGGGTGGCTGAGGGCTCAGGCTGCTTGGGGTGGACCTCTCTCTGGGC
CGCCAGGCTTTCAGCTCCATCCCTGCCCTCCAGTTTCCTCTCCATGGTCGCCACCAGGAT
CTTGAGCCACGTGTTCTCGGTAAGGAGGTTTTCTGAGGCATCGGCCAGCTCCTGCAGCTT
GCGGGCCAGCTCCTGCAGCTCACGTGTCTTCTGCTCATA

FIG. 28

BLASTP - alignment of 59_PROTEIN (SEQ ID NO:2) against swiss|Q10057|PDI1_SCHPO (SEQ ID NO:3)

This hit is scoring at : 7e-10 (expectation value), Alignment length (overlap): 131, Identities : 32 %, Scoring matrix : BLOSUM62 (used to infer consensus pattern), Database searched : nrdb

```
Q:    33 SQHLITEVTTDTFWEVVLQK-QDVLLLYYAPWCGFCPSLNHIFIQLARNLPMDT-FTVAR
         SQ. :. :..D.F ::V:.: :DVL: :YAPWCG.C.:L  .: :LA.... D: ..VA:
H:   353 SQEDLVVLVADNFDDIVMDETKDVLVEFYAPWCGHCKNLAPTYEKLAEEYSDDSNVVVAK
```

disulfe redox-active site CXXC motif

BLOCKS for Thioredoxin family proteins

```
         IDVSQNDLPWEFMVDRLPTVLFFPCNRKDLSVKYPEDVPITLPNLLRFILHHS--DPASS
         ID.::ND:  ...:. .PT::FF..N K  .V:Y. D   TL.:L .FI .H:  :P...
         IDATENDI--SVSISGFPTIMFFKANDKVNPVRYEGD--RTLEDLSAFIDKHASFEPIKK

         PQNVANSPTKE    159
         .:. . :P. E
         EKESVPAPDLE    479
```

FIG. 29

BLOCKS search results

AC#	Description	Strength	Score
BL00194	Thioredoxin family proteins.	1242	1223
AA# 55	LLLYYAPWCGfCP		

FIG. 30

HMMPFAM - alignment of 59_PROTEIN (SEQ ID NO:2) against pfam|hmm|thiored (SEQ ID NO:4)

Thioredoxin
This hit is scoring at : 26.4; Expect = 1.2e-06, Scoring matrix : BLOSUM62 (used to infer consensus pattern),

```
Q:      34 QHLITEVTTDTFWEVVLQ---KQDVLLLYYAPWCGFC      67
           ..::. :T.:.F E VL:    .:.VL: :YAPWCG C
H:       1 ssvvvvltdenFdeevlkaksdkpVLVdFyApWCGpC      37
```

Fig. 31

atgaagctcaccctagagtcttttattcaaaacttcagcgttctctatagtcccttgaaa
aggcatctcattggaagtggctctgcccagttcccgtctcagcatttaatcactgaagtg
acaactgatacctttgggaagtagtccttcaaaaacaggacgttctcctgctctattac
gctccgtggtgcggcttctgtccatccctcaatcacatcttcatccagctagctcggaac
ctgcccatggacacattcactgtggcaaggattgacgtgtctcagaatgaccttccttgg
gaatttatggtcgatcgtcttcctactgtcttgtttttccctgcaacag

Fig. 32

MKLTLESFIQ NFSVLYSPLK RHLIGSGSAQ FPSQHLITEV TTDTFWEVVL QKQDVLLLYY
APWCGFCPSL NHIFIQLARN LPMDTFTVAR IDVSQNDLPW EFMVDRLPTV LFFPCN

Fig. 33

BLASTP - alignment of 59_short_TR1 against swiss|P30101|ER60_HUMAN
This hit is scoring at : 4e-09 (expectation value), Alignment length (overlap) : 116, Identities : 33 %, Scoring matrix : BLOSUM62 (used to infer consensus pattern), Database searched : nrdb

```
Q:      5 LESFIQNFSVLYSPLKRHLIGSGSAQFPSQH--LITEVTTDTFWEVVL-QKQDVLLLYYA
          LE.F:Q::  . . LKR:L    S...P..:    :. V..:.F E:V  :.:DVL: :YA
H:    349 LERFLQDY--FDGNLKRYL---KSEPIPESNDGPVKVVVAENFDEIVNNENKDVLIEFYA

          PWCGFCPSLNHIFIQLARNLPMD-TFTVARIDVSQNDLPWEFMVDRLPTVLFFPCN     116
          PWCG.C.:L.  : :L...L..D ...:A::D.:.ND:P .: V  .PT:.F P.N
          PWCGHCKNLEPKYKELGEKLSKDPNIVIAKMDATANDVPSPYEVRGFPTIYFSPAN     459
```

disulfe redox-active site CXXC motif

BLOCKS for Thioredoxin family proteins

BLOCKS search results

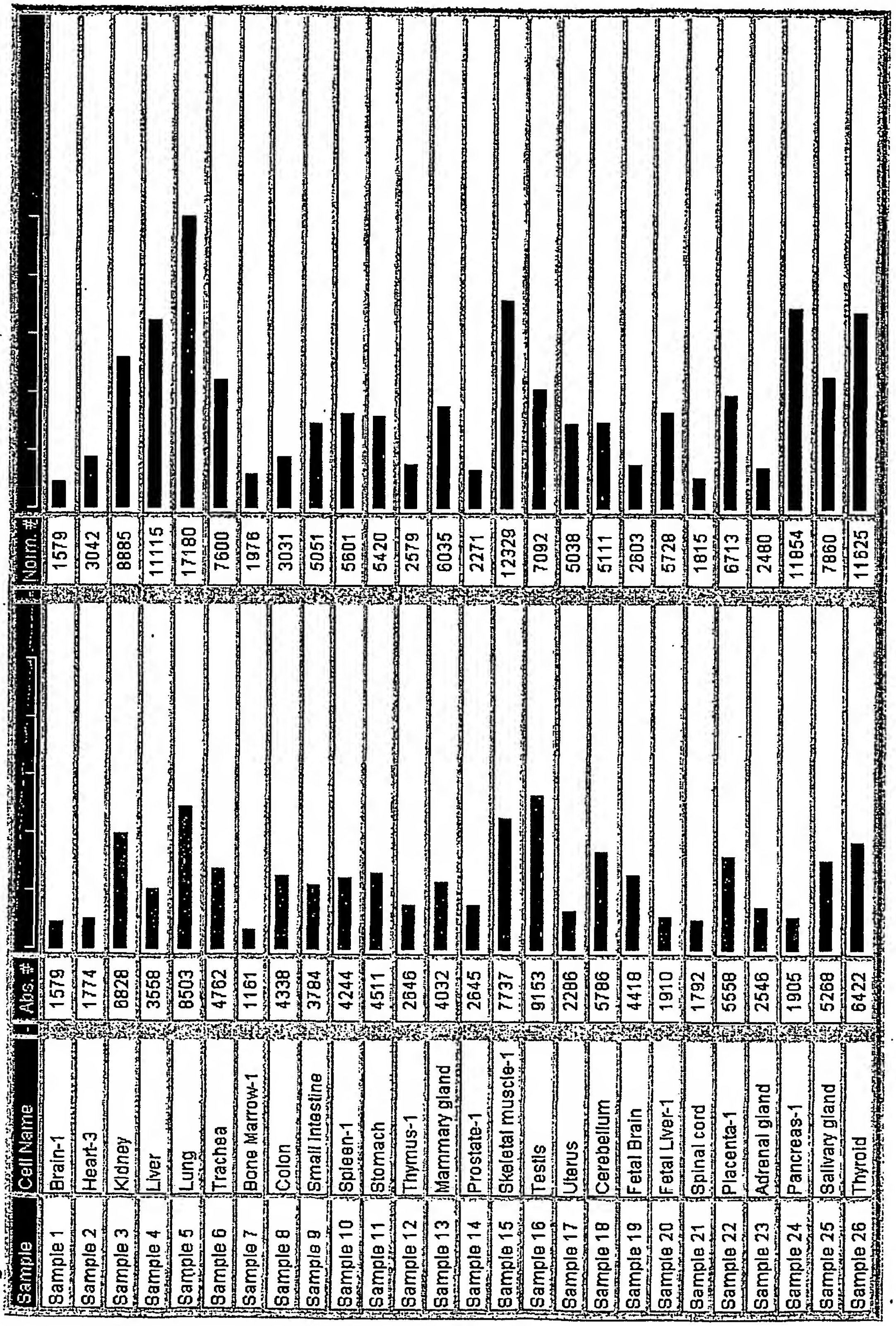
AC#	Description	Strength	Score
BL00194	Thioredoxin family proteins.	1242	1223
AA# 55	LLLYYAPWCGfCP		

Fig. 34

HMMPFAM - alignment of 59_short_TR1 against pfam|hmm|thiored Thioredoxin, This hit is scoring at : 26.4; Expect = 1.2e-06, Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:      34 QHLITEVTTDTFWEVVLQ---KQDVLLLYYAPWCGFC      67
              ..::. :T.:.F E VL:    .:.VL: :YAPWCG C
H:       1 ssvvvvltdenFdeevlkaksdkpVLVdFyApWCGpC      37
```

Fig. 35



Sample	Cell Name	Abs. #	Norm. #
Sample 1	Brain-1	1579	1579
Sample 2	Heart-3	1774	3042
Sample 3	Kidney	6828	8885
Sample 4	Liver	3558	11115
Sample 5	Lung	8503	17180
Sample 6	Trachea	4762	7600
Sample 7	Bone Marrow-1	1161	1976
Sample 8	Colon	4338	3031
Sample 9	Small Intestine	3784	5051
Sample 10	Spleen-1	4244	5601
Sample 11	Stomach	4511	5420
Sample 12	Thymus-1	2646	2579
Sample 13	Mammary gland	4032	6035
Sample 14	Prostate-1	2645	2271
Sample 15	Skeletal muscle-1	7737	12329
Sample 16	Testis	9153	7092
Sample 17	Uterus	2286	5038
Sample 18	Cerebellum	5786	5111
Sample 19	Fetal Brain	4418	2603
Sample 20	Fetal Liver-1	1910	5728
Sample 21	Spinal cord	1792	1815
Sample 22	Placenta-1	5558	6713
Sample 23	Adrenal gland	2546	2480
Sample 24	Pancreas-1	1905	11854
Sample 25	Salivary gland	5268	7860
Sample 26	Thyroid	6422	11625

Fig. 36

Sample	Cell Name	Abs. #	Norm. #
Sample 1	lymph node	285.9	285
Sample 2	PBL	453.3	330
Sample 3	Tonsil	107	545
Sample 4	PBMN	785.9	247
Sample 5	PBMN stimulated	498.1	417
Sample 6	T-cell(CD8+)	4814	447
Sample 7	T-cell(CD8+)activated	142.8	284
Sample 8	T-cell(CD4+)	3715	443
Sample 9	T-cell(CD4+) activated	568.4	1453
Sample 10	B-cell(CD19+)	2282	735
Sample 11	B-cell(CD19+) activated	354.5	790
Sample 12	Monocytes(CD14+)	510.8	201
Sample 13	Th1 clone	2781	330
Sample 14	Th2 clone	1863	78
Sample 15	neutroA-2	995.7	24
Sample 16	neutroB-2	547.8	23
Sample 17	Normal Bronchial/Tracheal Epithelial Cells	539	2548
Sample 18	Normal Bronchial/Tracheal smooth muscle cell	423.2	1381
Sample 19	Normal lung fibroblast	863.7	587
Sample 20	Microvascular Endothelial cell	1863	1566
Sample 21	U937	1672	3694
Sample 22	RAMOS	498	2205
Sample 23	Jurkat	751.4	2385
Sample 24	HelaS3	966.7	1228
Sample 25	IMR-90	609.2	767
Sample 26	HEK293	54.35	26

SEQUENCE LISTING

<110> Bayer AG

Bull, Christof

<120> REGULATION OF HUMAN PROTEIN DISULFIDE ISOMERASE-LIKE ENZYME

<130> Lio132 foreign countries

<150> 60/231,297

<151> 2000-09-08

<160> 29

<170> PatentIn version 3.1

<210> 1

<211> 993

<212> DNA

<213> Homo sapiens

<400> 1

```
atgaagctca ccctagagtc ttttattcaa aacttcagcg ttctctatag tcccttgaaa      60
aggcatctca ttggaagtgg ctctgcccag ttcccgtctc agcatttaat cactgaagtg     120
acaactgata ccttttggga agtagtcctt caaaaacagg acgttctcct gctctattac     180
gctccgtggt gcggcttctg tccatccctc aatcacatct tcatccagct agctcggaac     240
ctgcccatgg acacattcac tgtggcaagg attgacgtgt ctcagaatga ccttccttgg     300
gaatttatgg tcgatcgtct tcctactgtc ttgtttttc cctgcaacag aaaggaccta     360
agtgtgaaat accccgaaga cgtccctatc acccttccaa acctgttgag gttcattttg     420
catcactcag accctgcttc cagcccccag aatgtggcta actctcctac caaggagtgt     480
cttcagagcg aggcagtctt acagcggggg cacatctccc acttggagag agagatccag     540
aaactgagag cagaaataag cagcctccag cgagcacaag tgcaggtgga gtcccagctc     600
tccagtgccc gcagagatga gcaccggctg cggcagcagc agcgggccct ggaagagcag     660
cacagcctgc tccacgcaca cagtgagcag ctgcaggccc tctatgagca gaagacacgt     720
gagctgcagg agctggcccg caagctgcag gagctggccg atgcctcaga aaacctcctt     780
accgagaaca cgtggctcaa gatcctggtg gcgaccatgg agaggaaact ggagggcagg     840
gatggagctg aaagcctggc ggcccagaga gaggtccacc ccaagcagcc tgagccctca     900
```

gccaccccccc agctccctgg cagctcccct ccacctgcca atgtcagcgc cacactggtg 960
tctgaaagga ataaggagaa caggacagac taa 993

<210> 2

<211> 330

<212> PRT

<213> Homo sapiens

<400> 2

Met Lys Leu Thr Leu Glu Ser Phe Ile Gln Asn Phe Ser Val Leu Tyr
1 5 10 15
Ser Pro Leu Lys Arg His Leu Ile Gly Ser Gly Ser Ala Gln Phe Pro
20 25 30
Ser Gln His Leu Ile Thr Glu Val Thr Thr Asp Thr Phe Trp Glu Val
35 40 45
Val Leu Gln Lys Gln Asp Val Leu Leu Leu Tyr Tyr Ala Pro Trp Cys
50 55 60
Gly Phe Cys Pro Ser Leu Asn His Ile Phe Ile Gln Leu Ala Arg Asn
65 70 75 80
Leu Pro Met Asp Thr Phe Thr Val Ala Arg Ile Asp Val Ser Gln Asn
85 90 95
Asp Leu Pro Trp Glu Phe Met Val Asp Arg Leu Pro Thr Val Leu Phe
100 105 110
Phe Pro Cys Asn Arg Lys Asp Leu Ser Val Lys Tyr Pro Glu Asp Val
115 120 125
Pro Ile Thr Leu Pro Asn Leu Leu Arg Phe Ile Leu His His Ser Asp
130 135 140
Pro Ala Ser Ser Pro Gln Asn Val Ala Asn Ser Pro Thr Lys Glu Cys
145 150 155 160
Leu Gln Ser Glu Ala Val Leu Gln Arg Gly His Ile Ser His Leu Glu
165 170 175
Arg Glu Ile Gln Lys Leu Arg Ala Glu Ile Ser Ser Leu Gln Arg Ala
180 185 190
Gln Val Gln Val Glu Ser Gln Leu Ser Ser Ala Arg Arg Asp Glu His
195 200 205
Arg Leu Arg Gln Gln Gln Arg Ala Leu Glu Glu Gln His Ser Leu Leu
210 215 220
His Ala His Ser Glu Gln Leu Gln Ala Leu Tyr Glu Gln Lys Thr Arg
225 230 235 240
Glu Leu Gln Glu Leu Ala Arg Lys Leu Gln Glu Leu Ala Asp Ala Ser
245 250 255
Glu Asn Leu Leu Thr Glu Asn Thr Trp Leu Lys Ile Leu Val Ala Thr
260 265 270
Met Glu Arg Lys Leu Glu Gly Arg Asp Gly Ala Glu Ser Leu Ala Ala
275 280 285
Gln Arg Glu Val His Pro Lys Gln Pro Glu Pro Ser Ala Thr Pro Gln
290 295 300

Leu Pro Gly Ser Ser Pro Pro Pro Ala Asn Val Ser Ala Thr Leu Val
305 310 315 320

Ser Glu Arg Asn Lys Glu Asn Arg Thr Asp
325 330

<210> 3

<211> 492

<212> PRT

<213> Homo sapiens

<400> 3

Met Lys Ile Ser Asn Leu Leu Ala Ala Phe Leu Ala Phe Ser Gly Gly
1 5 10 15

Phe Phe Cys Ala Ser Ala Glu Val Pro Lys Val Asn Lys Glu Gly Leu
20 25 30

Asn Glu Leu Ile Thr Ala Asp Lys Val Leu Met Val Lys Phe Tyr Ala
35 40 45

Pro Trp Cys Gly His Cys Lys Ala Leu Ala Pro Glu Tyr Glu Ser Ala
50 55 60

Ala Asp Glu Leu Glu Lys Asp Gly Ile Ser Leu Val Glu Val Asp Cys
65 70 75 80

Thr Glu Glu Gly Asp Leu Cys Ser Glu Tyr Ser Ile Arg Gly Tyr Pro
85 90 95

Thr Leu Asn Val Phe Lys Asn Gly Lys Gln Ile Ser Gln Tyr Ser Gly
100 105 110

Pro Arg Lys His Asp Ala Leu Val Lys Tyr Met Arg Lys Gln Leu Leu
115 120 125

Pro Thr Val Lys Pro Ile Ser Lys Asp Thr Leu Glu Asn Phe Val Glu
130 135 140

Lys Ala Asp Asp Leu Ala Val Val Ala Phe Phe Lys Asp Gln Lys Leu
145 150 155 160

Asn Asp Thr Tyr Thr Glu Val Ala Glu Val Met Lys Asp Asp Phe Val
165 170 175

Phe Ala Ala Ser Asp Asp Lys Glu Leu Ala Lys Ser Leu Gly Ser Asn
180 185 190

Phe Pro Gly Ile Val Ala Phe Thr Lys Asp Ala Ala Gln Asp Ser Asp
195 200 205

Lys Leu Val Tyr Thr Gly Asp Trp Asp Pro Ala Ser Ile Ala Asp Phe
210 215 220

Ile Gly Val Ser Ser Ile Pro Leu Leu Asp Glu Leu Asn Gln Met Thr
225 230 235 240

Phe Gly Lys Tyr Gln Gln Ser Gly Leu Pro Leu Gly Ile Ile Phe Tyr
245 250 255

Asn Ser Thr Glu Ser Arg Asp Glu Leu Tyr Asp Val Phe Gln Pro Leu
260 265 270

Ala Lys Lys Tyr Gln Asp Thr Leu Arg Phe Ala Phe Leu Asp Ala Val
275 280 285

Arg Tyr Gly Ala Val Ala Lys Gln Met Asn Val Glu Ser Asp Trp Pro
290 295 300

Ala Phe Val Ile Ala Asn Leu Lys Ser Met Leu Lys Tyr Pro Phe Pro
305 310 315 320

Thr Thr Glu Leu Thr Ala Lys Ala Met Thr Lys Phe Val Gly Asp Phe
325 330 335

Val Asp Gly Lys Leu Gln Pro Lys Ile Lys Ser Gln Pro Ile Pro Glu
340 345 350

Ser Gln Glu Asp Leu Val Val Leu Val Ala Asp Asn Phe Asp Asp Ile
355 360 365

Val Met Asp Glu Thr Lys Asp Val Leu Val Glu Phe Tyr Ala Pro Trp
370 375 380

Cys Gly His Cys Lys Asn Leu Ala Pro Thr Tyr Glu Lys Leu Ala Glu
385 390 395 400

Glu Tyr Ser Asp Asp Ser Asn Val Val Val Ala Lys Ile Asp Ala Thr
405 410 415

Glu Asn Asp Ile Ser Val Ser Ile Ser Gly Phe Pro Thr Ile Met Phe
420 425 430

Phe Lys Ala Asn Asp Lys Val Asn Pro Val Arg Tyr Glu Gly Asp Arg
435 440 445

Thr Leu Glu Asp Leu Ser Ala Phe Ile Asp Lys His Ala Ser Phe Glu
450 455 460

Pro Ile Lys Lys Glu Lys Glu Ser Val Pro Ala Pro Asp Leu Glu Asp
465 470 475 480

Gln Val Ala Val Glu Asp Glu Met Ala Asp Glu Leu
485 490

<210> 4

<211> 37

<212> PRT

<213> Homo sapiens

<400> 4

Ser Ser Val Val Val Val Leu Thr Asp Glu Asn Phe Asp Glu Glu Val
1 5 10 15

Leu Lys Ala Lys Ser Asp Lys Pro Val Leu Val Asp Phe Tyr Ala Pro
20 25 30

Trp Cys Gly Pro Cys
35

<210> 5

<211> 1532

<212> DNA

<213> Homo sapiens

<400> 5
ttttaacttc ctatagcccc ttcagctact acactgcatg ttgcaggacc ataagcaggg 60
gtgtgtcagg cttcatcgac tctgaacaag gtgtctttga agcccctact gttgcatttt 120
cttcccttga gaagaaatgt gaggttgatg ccccaagctc cgttcctcac attgaggaga 180
acaggtatct ctttccagaa gtggacatga ctagcacaaa cttcacaggc ctgagctgca 240
gaaccaacaa gactctcaac atctaccttt tggattcaaa tttgtttggt tatatgcaga 300
gagactgggt gctccgagct ccactcaggt gaaagaattt gcggcaattg ttgacgtgaa 360
agaagaatct cattacatct tggatccaaa gcaagcactg atgaagctca ccctagagtc 420
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg 480
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga 540
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg 600
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac 660
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct 720
tcctactgtc ttgttttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga 780
cgtccctatc acccttccaa acctgttgag gttcattttg catcactcag accctgcttc 840
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt 900
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag 960
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga 1020
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca 1080
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg 1140
caagctgcag gagctggccg atgcctcaga aaacctcctt accgagaaca cgtggctcaa 1200
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc 1260
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg 1320
cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa 1380
caggacagac taacttttta aatgatatga agaaatcaga ggtgaaaatt gtacattggg 1440
aatatattta tgcaaatttt attgaaattt attgtaaata aagattttct cagtggtcta 1500
gaaaatcaaa aaaaaaaaaa aaaaaaaaaa aa 1532

<210> 6

<211> 1532

<212> DNA

<213> Homo sapiens

<400> 6
ttttaacttc ctatagcccc ttcagctact acactgcatg ttgcaggacc ataagcaggg 60
gtgtgtcagg cttcatcgac tctgaacaag gtgtctttga agcccctact gttgcatttt 120
cttcccttga gaagaaatgt gaggttgatg ccccaagctc cgttcctcac attgaggaga 180
acaggtatct ctttccagaa gtggacatga ctagcacaaa cttcacaggc ctgagctgca 240
gaaccaacaa gactctcaac atctaccttt tggattcaaa tttgtttggt tatatgcaga 300
gagactgggt gctccgagct ccactcaggt gaaagaattt gcggcaattg ttgacgtgaa 360
agaagaatct cattacatct tggatccaaa gcaagcactg atgaagctca ccctagagtc 420
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg 480
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga 540
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg 600
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac 660

```
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct    720
tcctactgtc ttgtttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga    780
cgtccctatc acccttccaa acctgttgag gttcattttg catcactcag accctgcttc    840
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt    900
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag    960
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga   1020
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca   1080
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg   1140
caagctgcag gagctggccg atgcctcaga aacctcctt accgagaaca cgtggctcaa   1200
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc   1260
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg   1320
cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa   1380
caggacagac taacttttta aatgatatga agaaatcaga ggtgaaaatt gtacattggg   1440
aatatattta tgcaaatttt attgaaattt attgtaaata aagattttct cagtggtcta   1500
gaaaatcaaa aaaaaaaaaa aaaaaaaaaa aa                                 1532
```

<210> 7

<211> 1532

<212> DNA

<213> Homo sapiens

<400> 7

```
ttttaacttc ctatagcccc ttcagctact acactgcatg ttgcaggacc ataagcaggg     60
gtgtgtcagg cttcatcgac tctgaacaag gtgtctttga agcccctact gttgcatttt    120
cttcccttga gaagaaatgt gaggttgatg ccccaagctc cgttcctcac attgaggaga    180
acaggtatct ctttccagaa gtggacatga ctagcacaaa cttcacaggc ctgagctgca    240
gaaccaacaa gactctcaac atctaccttt tggattcaaa tttgtttggt tatatgcaga    300
gagactgggt gctccgagct ccactcaggt gaaagaattt gcggcaattg ttgacgtgaa    360
agaagaatct cattacatct tggatccaaa gcaagcactg atgaagctca ccctagagtc    420
ttttattcaa aacttcagcg ttctctatag tcccttgaaa aggcatctca ttggaagtgg    480
ctctgcccag ttcccgtctc agcatttaat cactgaagtg acaactgata ccttttggga    540
agtagtcctt caaaaacagg acgttctcct gctctattac gctccgtggt gcggcttctg    600
tccatccctc aatcacatct tcatccagct agctcggaac ctgcccatgg acacattcac    660
tgtggcaagg attgacgtgt ctcagaatga ccttccttgg gaatttatgg tcgatcgtct    720
tcctactgtc ttgtttttc cctgcaacag aaaggaccta agtgtgaaat accccgaaga    780
cgtccctatc acccttccaa acctgttgag gttcattttg catcactcag accctgcttc    840
cagcccccag aatgtggcta actctcctac caaggagtgt cttcagagcg aggcagtctt    900
acagcggggg cacatctccc acttggagag agagatccag aaactgagag cagaaataag    960
cagcctccag cgagcacaag tgcaggtgga gtcccagctc tccagtgccc gcagagatga   1020
gcaccggctg cggcagcagc agcgggccct ggaagagcag cacagcctgc tccacgcaca   1080
cagtgagcag ctgcaggccc tctatgagca gaagacacgt gagctgcagg agctggcccg   1140
caagctgcag gagctggccg atgcctcaga aacctcctt accgagaaca cgtggctcaa   1200
gatcctggtg gcgaccatgg agaggaaact ggagggcagg gatggagctg aaagcctggc   1260
ggcccagaga gaggtccacc ccaagcagcc tgagccctca gccacccccc agctccctgg   1320
cagctcccct ccacctgcca atgtcagcgc cacactggtg tctgaaagga ataaggagaa   1380
caggacagac taacttttta aatgatatga agaaatcaga ggtgaaaatt gtacattggg   1440
aatatattta tgcaaatttt attgaaattt attgtaaata aagattttct cagtggtcta   1500
```

gaaaatcaaa aaaaaaaaaa aaaaaaaaaa aa 1532

<210> 8

<211> 759

<212> DNA

<213> Homo sapiens

<400> 8

accactgaga aaatctttat ttacaataaa tttcaataaa atttgcataa atatattccc 60
aatgtacaat tttcacctct gatttcttca tatcatttaa aaagttagtc tgtcctgttc 120
tccttattcc tttcagacac cagtgtggcg ctgacattgg caggtggagg ggagctgcca 180
gggagctggg gggtggctga gggctcaggc tgcttggggt ggacctctct ctgggccgcc 240
aggctttcag ctccatccct gccctccagt ttcctctcca tggtcgccac caggatcttg 300
agccacgtgt tctcggtaag gaggttttct gaggcatcgg ccagctcctg cagcttgcgg 360
gccagctcct gcagctcacg tgtcttctgc tcatagaggg cctgcagctg ctcactgtgt 420
gcgtggagca ggctgtgctg ctcttccagg gcccgctgct gctgccgcag ccggtgctca 480
tctctgcggg cactggagag ctgggactcc acctgcactt gtgctcgctg gaggctgctt 540
atttctgctc tcagtttctg gatctctctc tccaagtggg agatgtgccc ccgctgtaag 600
actgcctcgc tctgaagaca ctccttggta ggagagttag ccacattctg gggctggaa 660
agcagggtct gagtgatgca aaatgaacct caacaggttt ggaagggtga tggggacgtc 720
ttcggggtat ttcacactta gggtcctttc tgttgcagg 759

<210> 9

<211> 775

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (653)..(653)

<223> n = a,t,c, or g

<400> 9

tagaccactg agaaaatctt tatttacaat aaatttcaat aaaatttgca taaatatatt 60
cccaatgtac aattttcacc tctgatttct tcatatcatt taaaaagtta gtctgtcctg 120
ttctccttat tcctttcaga caccagtgtg gcgctgacat tggcaggtgg aggggagctg 180
ccagggagct gggggtggc tgagggctca ggctgcttgg ggtggacctc tctctgggcc 240
gccaggcttt cagctccatc cctgccctcc agtttcctct ccatggtcgc caccaggatc 300
ttgagccacg tgttctcggt aaggaggttt tctgaggcat cggccagctc ctgcagcttg 360
cgggccagct cctgcagctc acgtgtcttc tgctcataga gggcctgcag ctgctcactg 420

tgtgcgtgga gcaagctgtg ctgctcttcc agggcccgct gctgctgccg cagccggtgc 480
tcatctctgc gggcactgga gagctgggac tccacctgca cttgtgctcg ctggaggctg 540
cttatttctg ctctcagttt ctggatctct ctctccaagt gggagatgtg cccccgctgt 600
aagactgcct cgctctgaag acactccttg gtaggagagt tagccacatt ctnggggctg 660
gaagcagggt ctgaatgatg caaaatgaac cctaacaggt ttggaaaggt gatggggagg 720
tcttccgggg tatttcacac ttaggtccct tctgttggac ggaaaaaaaa cagcg 775

<210> 10

<211> 533

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (211)..(211)

<223> n = a,t,g, or c

<400> 10
tagaccactg agaaaatctt tatttacaat aaatttcaat aaaatttgca taaatatatt 60
cccaatgtac aattttcacc tctgatttct tcatatcatt taaaaagtta gtctgtcctg 120
ttctccttat tcctttcaga caccagtgtg gcgctgacat tggcaggtgg aggggagctg 180
ccagggagct ggggggtggc tgagggctca ngctgcttgg ggtggacctc tctctgggcc 240
gccaggcttt cagctccatc cctgccctcc agtttcctct ccatggtcgc caccaggatc 300
ttgagccacg tgttctcggt aaggaggttt tctgaggcat cggccagctc ctgcagcttg 360
cgggccagct cctgcagctc acgtgtcttc tgctcataga gggcctgcag ctgctcactg 420
tgtgcgtgga gcaggctgtg ctgctcttcc agggcccgct gctgctgccg cagccggtgc 480
tcatctctgc gggcactgga gagctgggac tccacctgca cttgtgctcg ctg 533

<210> 11

<211> 543

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (371)..(371)

<223> n = a,t,g, or c

<220>

<221> misc_feature

<222> (407)..(407)

<223> n = a,t,g, or c

<400> 11
cgagctccac tcaggtgaaa gaatttgcgg caattgttga cgtgaaagaa gaatctcatt 60
acatcttgga tccaaagcaa gcactgatga agctcacccct agagtctttt attcaaaact 120
tcagcgttct ctatagtccc ttgaaaaggc atctcattgg aagtggctct gcccagttcc 180
cgtctcagca tttaatcact gaagtgacaa ctgatacctt ttgggaagta gtccttcaaa 240
aacaggacgt tctcctgctc tattacgctc cgtggtgcgg cttctgtcca tccctcaatc 300
acatcttcat ccagctagct cggaacctgc ccatggacac attcactgtg gcaaggattg 360
acgtgtctca naatgacctt ccttgggaat ttatggtcga tcgtctncct actgtcttgt 420
tttttccctg caacagaaag gacctaagtg tgaaataccc cgaagacgtc ctcatcaccc 480
ttccaaacct ggtgaggttc attttgcatc actcagaccc tgcttccagc ccccagatgt 540
ggc 543
<210> 12

<211> 491

<212> DNA

<213> Homo sapiens

<400> 12
tagaccactg agaaaatctt tatttacaat aaatttcaat aaaatttgca taaatatatt 60
cccaatgtac aattttcacc tctgatttct tcatatcatt taaaaagtta gtctgtcctg 120
ttctccttat tcctttcaga caccagtgtg gcgctgacat tggcaggtgg aggggagctg 180
ccagggagct gggggtggc tgagggctca ggctgcttgg ggtggacctc tctctgggcc 240
gccaggcttt cagctccatc cctgccctcc agtttcctct ccatggtcgc caccaggatc 300
ttgagccacg tgttctcggt aaggaggttt tctgaggcat cggccagctc ctgcagcttg 360
cgggccagct cctgcagctc acgtgtcttc tgctcataga gggcctgcag ctgctcactg 420
tgtgcgtgga gcaggctgtg ctgctcttcc agggcccgct gctgctgccg cagccggtgc 480
tcatctctgc g 491
<210> 13

<211> 488

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (404)..(404)

<223> n=a,t,g, or c

<220>

<221> misc_feature

<222> (406)..(406)

<223> n=a,t,g, or c

<220>

<221> misc_feature

<222> (455)..(455)

<223> n=a,t,g, or c

<400> 13
ctgaagtgac aactgatacc ttttgggaag tagtccttca aaaacaggac gttctcctgc 60
tctattacgc tccgtggtgc ggcttctgtc catccctcaa tcacatcttc atccagctag 120
ctcggaacct gcccatggac acattcactg tggcaaggat tgacgtgtct cagaatgacc 180
ttccttggga atttatggtc gatcgtcttc ctactgtctt gttttttccc tgcaacagaa 240
aggacctaag tgtgaaatac cccgaagacg tccccatcac ccttccaaac ctgttgaggt 300
tcattttgca tcactcagac cctgcttcca gcccccagaa tgtggctaac tctcctacca 360
aggagtgtct tcagagcgag gcagtcttac agcgggggca catntnccac ttggagagag 420
agatccagaa actgagagca gaaataagca gcctncagcg agcacaagtg caggtggagt 480
cccagctc 488

<210> 14

<211> 468

<212> DNA

<213> Homo sapiens

<400> 14
ctagaccact gagaaaatct ttatttacaa taaatttcaa taaaatttgc ataaatatat 60
tcccaatgta caattttcac ctctgatttc ttcatatcat ttaaaaagtt agtctgtcct 120
gttctcctta ttcctttcag acaccagtgt ggcgctgaca ttggcaggtg gaggggagct 180
gccagggagc tgggggggtgg ctgagggctc aggctgcttg ggtggacct ctctctgggc 240
cgccaggctt tcagctccat ccctgccctc cagtttcctc tccatggtcg ccaccaggat 300

cttgagccac gtgttctcgg taaggaggtt ttctgaggca tcggccagct cctgcagctt 360
gcgggccagc tcctgcagct cacgtgtctt ctgctcatag agggcctgca gctgctcact 420
gtgtgcgtgg agcaggctgt gctgctcttc cagggcccgc tgctgctg 468

<210> 15

<211> 459

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (7)..(7)

<223> n = a,t,g, or c

<400> 15
ttttttnatg tttttttag accactgaga aaatctttat ttacaataaa tttcaataaa 60
atttgcataa atatattccc aatgtacaat tttcacctct gatttcttca tatcatttaa 120
aaagttagtc tgtcctgttc tccttattcc tttcagacac cagtgtggcg ctgacattgg 180
caggtggagg ggagctgcca gggagctggg gggtggctga gggctcaggc tgcttggggt 240
ggacctctct ctgggccgcc aggctttcag ctccatccct gccctccagt ttcctctcca 300
tggtcgccac caggatcttg agccacgtgt tctcggtaag gaggttttct gaggcatcgg 360
ccagctcctg cagcttgcgg gccagctcct gcagctcacg tgtcttctgc tcatagaggg 420
cctgcagctg ctcactgtgt gcgtggagca ggctgtgct 459

<210> 16

<211> 472

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (328)..(329)

<223> n = a,t,g, or c

<400> 16
tctagaccac tgagaaaatc tttatttaca ataaatttca ataaaatttg cataaatata 60
ttcccaatgt acaattttca cctctgattt cttcatatca tttaaaaagt tagtctgtcc 120

tgttctcctt attcctttca gacaccagtg tggcgctgac attggcaggt ggaggggagc 180
tgccagggag ctggggggtg gctgagggct caggctgctt ggggtggacc tctctctggg 240
ccgccacggt ttcagctcca tccctgccct ccagtttcct ctccatggtc gccaccagga 300
tcttgagcca cgtgttctcc gtaagagnnt tttctgaggc atcggccagc tcctgcagct 360
tgcgagccag ctcctgcagc tcacgtgtct tctgctcata gagggcctgc agctgctcac 420
tgtgtgcgtg gagcaggctg tgctgctctt ccagggcccg ctgctgctgc cg 472

<210> 17

<211> 434

<212> DNA

<213> Homo sapiens

<400> 17
tctagaccac tgagaaaatc tttatttaca ataaatttca ataaaatttg cataaatata 60
ttcccaatgt acaattttca cctctgattt cttcatatca tttaaaaagt tagtctgtcc 120
tgttctcctt attcctttca gacaccagtg tggcgctgac attggcaggt ggaggggagc 180
tgccagggag ctggggggtg gctgagggct caggctgctt ggggtggacc tctctctggg 240
ccgccaggct ttcagctcca tccctgccct ccagtttcct ctccatggtc gccaccagga 300
tcttgagcca cgtgttctcg gtaaggaggt tttctgaggc atcggccagc tcctgcagct 360
tgcgggccag ctcctgcagc tcacgtgtct tctgctcata gagggcctgc agctgctcac 420
tgtgtgcgtg gagc 434

<210> 18

<211> 440

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (218)..(218)

<223> n = a,t,g, or c

<220>

<221> misc_feature

<222> (227)..(227)

<223> n = a,t,g, or c

<220>

<221> misc_feature

<222> (237)..(237)

<223> n = a,t,g, or c

<400> 18
gattttctag accactgaga aaatctttat ttacaataaa tttcaataaa atttgcataa 60
atatattccc aatgtacaat tttcacctct gatttcttca tatcatttaa aaagttagtc 120
tgtcctgttc tccttattcc tttcagacac cagtgtggcg ctgacattgg caggtggagg 180
ggagctgcca gggagctggg gggtggctga gggctcangc tgcttgnggt ggacctntct 240
ctgggccgcc aggctttcag ctccatccct gccctccagt ttcctctcca tggtcgccac 300
caggatcttg agccacgtgt tctcggtaag gaggttttct gaggcatcgg ccagctcctg 360
cagcttgcgg gccagctcct gcagctcacg tgtcttctgc tcatagaggg cctgcagctg 420
ctcactgtgt gcgtggagca 440
<210> 19

<211> 446

<212> DNA

<213> Homo sapiens

<400> 19
agaccattta gtaaatcttt atttacaata aatttcaatg aaatttgcat aaatatattc 60
ccaatgtaca attttcacct ctgatttctt catatcattt aaaaagttag tctgtcctgt 120
tctccttatt cctttcagac accagtgtgg cgctgacatt ggcaggtgga ggggagctgc 180
cagggagctg gggggtggct gagggctcag gctgcttggg gtggacctct ctctgggccg 240
ccaggctttc agctccatcc ctgccctcca gtttcctctc catggtcgcc accaggatct 300
tgagccacgt gttctcggta aggaggtttt ctgaggcatc ggccagctcc tgcagcttgc 360
gggccagctc ctgcagctca cgtgtcttct gctcatagag ggcctgcagc tgctcactgt 420
gtgcgtggag caggctgtgc tgctct 446
<210> 20

<211> 440

<212> DNA

<213> Homo sapiens

<400> 20
tttttttttc tagaccactg agaaaatctt tatttacaat aaatttcaat aaaatttgca 60
taaatatatt cccaatgtac aattttcacc tctgatttct tcatatcatt taaaaagtta 120
gtctgtcctg ttctccttat tcctttcaga caccagtgtg gcgctgacat tggcaggtgg 180
aggggagctg ccagggagct gggggtggc tgagggctca ggctgcttgg ggtggacctc 240

tctctgggcc gccacgtttc agctccatcc ctgccctcca gtttcctctc catggtcgcc 300
accaggatct tgagccacgt gttctcggta aggaggtttt ctgaggcatc ggccagctcc 360
tgcagcttgc gggccagctc ctgcagctca cgtgtcttct gctcatagag ggcctgcagc 420
tgctcactgt gtgcgtggag 440

<210> 21

<211> 441

<212> DNA

<213> Homo sapiens

<400> 21
tttttttttc tagaccactg agaaaatctt tatttacaat aaatttcaat aaaatttgca 60
taaatatatt cccaatgtac aattttcacc tctgatttct tcatatcatt taaaaagtta 120
gtctgtcctg ttctccttat tcctttcaga caccagtgtg gcgctgacat tggcaggtgg 180
aggggagctg ccagggagct gggggggtggc tgagggctca ggctgcttgg ggtggacctc 240
tctctgggcc gccacgttta cagctccatc cctgccctcc agtttcctct ccatggtcgc 300
caccaggatc ttgagccacg tgttctcggt aaggaggttt tctgaggcat cggccagctc 360
ctgcagcttg cgggccagct cctgcagctc acgtgtcttc tgctcataga gggcctgcag 420
ctgctcactg tgtgcgtgga g 441

<210> 22

<211> 437

<212> DNA

<213> Homo sapiens

<400> 22
ctagaccact gagaaaatct ttatttacaa taaatttcaa taaaatttgc ataaatatat 60
tcccaatgta caattttcac ctctgatttc ttcatatcat ttaaaaagtt agtctgtcct 120
gttctcctta ttcctttcag acaccagtgt ggcgctgaca ttggcaggtg gaggggagct 180
gccagggagc tgggggggtgg ctgagggctc aggctgcttg gggtggacct ctctctgggc 240
cgccacgttt cagctccatc cctgccctcc agtttcctct ccatggtcgc caccaggatc 300
ttgagccacg tgttctcggt aaggaggttt tctgaggcat cggccagctc ctgcagcttg 360
cgtgccagct cctgcagctc acgtgtcttc tgctcataga gggcctgcag ctgctcactg 420
tgtgcgtgga gcaggct 437

<210> 23

<211> 437

<212> DNA

<213> Homo sapiens

<400> 23

ttttctagac cactgagaaa atctttattt acaataaatt tcaataaaat ttgcataaat 60
atattcccaa tgtacaattt tcacctctga tttcttcata tcatttaaaa agttagtctg 120
tcctgttctc cttattcctt tcagacacca gtgtggcgct gacattggca ggtggagggg 180
agctgccagg gagctggggg gtggctgagg gctcaggctg cttggggtgg acctctctct 240
gggccgccac gtttatcagcc tccatccctg ccctccagtt tcctctccat ggtcgccacc 300
aggatcttga gccacgtgtt ctcggtaagg aggttttctg aggcatcggc cagctcctgc 360
agcttgcgag ccagctcctg cagctcacgt gtcttctgct catagagggc ctgcagctgc 420
tcactgtgtg cgtggag 437

<210> 24

<211> 441

<212> DNA

<213> Homo sapiens

<400> 24
ttttttctag accactgaga aaatctttat ttacaataaa tttcaataaa atttgcataa 60
atatattccc aatgtacaat tttcacctct gatttcttca tatcatttaa aaagttagtc 120
tgtcctgttc tccttattcc tttcagacac cagtgtggcg ctgacattgg caggtggacg 180
ggagcagcca gggagctggg gggtggctga gggctcaggc tgcttggggt ggacctctct 240
ctgggccgcc acgtttctag ctccatccct gccctccagt ttcctctcca tggtcgccac 300
caggatcttg agccacgtgt tctcggtaag gaggttttct gaggcatcgg ccagctcctg 360
cagcttgcgg gccagctcct gcagctcacg tgtcttctgc tcatagaggg gctgaagctg 420
ctcactgtgt gcgtggagaa g 441

<210> 25

<211> 414

<212> DNA

<213> Homo sapiens

<400> 25
tttttttttt tttttctaga ccactgagaa aatctttatt tacaataaat ttcaataaaa 60
tttgcataaa tatattccca atgtacaatt ttcacctctg atttcttcat atcatttaaa 120
aagttagtct gtcctgttct ccttattcct ttcagacacc agtgtggcgc tgacattggc 180
aggtggaggg gagctgccag ggagctgggg ggtggctgag ggctcaggct gcttggggtg 240
gacctctctc tgggccgcca ggctttcagc tccatccctg ccctccagtt tcctctccat 300
ggtcgccacc aggatcttga gccacgtgtt ctcggtaagg aggttttctg aggcatcggc 360
cagctcctgc agcttgcggg ccagctcctg cagctcacgt gtcttctgct cata 414

<210> 26

<211> 459

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (306)..(306)

<223> n = a,t,g, or c

<400> 26
ttttttttc tagaccactg agaaaatctt tatttacaat aaatttcaat aaaatttgca 60
taaatatatt cccaatgtac aattttcacc tctgatttct tcatatcatt taaaaagtta 120
gtctgtcctg ttctccttat tcctttcaga caccagtgtg gcgctgacat tggcaggtgg 180
aggggagctg ccagggagct gggggtggc tgagggctca ggctgcttgg ggtggacctc 240
tctctgggcc gccaggcttt cagctccatc cctgccctcc agtttcctct ccatggtcgc 300
caccangatc ttgagccacg tgttctcggt agggaggttt tctgaggcat cggccagctc 360
ctgcagcttg cgggccagct cctgcacttc aacgtgtctt ctgctcatta gagggcctgc 420
agctgctcac tgtgttgcgt ggagcaaggc ctgtgcttg 459
<210> 27

<211> 399

<212> DNA

<213> Homo sapiens

<400> 27
ctagaccact gagaaaatct ttatttacaa taaatttcaa taaaatttgc ataaatatat 60
tcccaatgta caattttcac ctctgatttc ttcatatcat ttaaaaagtt agtctgtcct 120
gttctcctta ttcctttcag acaccagtgt ggcgctgaca ttggcaggtg gaggggagct 180
gccagggagc tgggggtgg ctgagggctc aggctgcttg gggtggacct ctctctgggc 240
cgccaggctt tcagctccat ccctgccctc cagtttcctc tccatggtcg ccaccaggat 300
cttgagccac gtgttctcgg taaggaggtt ttctgaggca tcggccagct cctgcagctt 360
gcgggccagc tcctgcagct cacgtgtctt ctgctcata 399
<210> 28

<211> 350

<212> DNA

<213> Homo sapiens

<400> 28
atgaagctca ccctagagtc ttttattcaa aacttcagcg ttctctatag tcccttgaaa 60
aggcatctca ttggaagtgg ctctgcccag ttcccgtctc agcatttaat cactgaagtg 120

acaactgata ccttttggga agtagtcctt caaaaacagg acgttctcct gctctattac 180
gctccgtggt gcggcttctg tccatccctc aatcacatct tcatccagct agctcggaac 240
ctgcccatgg acacattcac tgtggcaagg attgacgtgt ctcagaatga ccttccttgg 300
gaatttatgg tcgatcgtct tcctactgtc ttgtttttc cctgcaacag 350

<210> 29

<211> 116

<212> PRT

<213> Homo sapiens

<400> 29

```
Met Lys Leu Thr Leu Glu Ser Phe Ile Gln Asn Phe Ser Val Leu Tyr
1               5                   10                  15
Ser Pro Leu Lys Arg His Leu Ile Gly Ser Gly Ser Ala Gln Phe Pro
            20                  25                  30
Ser Gln His Leu Ile Thr Glu Val Thr Thr Asp Thr Phe Trp Glu Val
        35                  40                  45
Val Leu Gln Lys Gln Asp Val Leu Leu Leu Tyr Tyr Ala Pro Trp Cys
    50                  55                  60
Gly Phe Cys Pro Ser Leu Asn His Ile Phe Ile Gln Leu Ala Arg Asn
65                  70                  75                  80
Leu Pro Met Asp Thr Phe Thr Val Ala Arg Ile Asp Val Ser Gln Asn
                85                  90                  95
Asp Leu Pro Trp Glu Phe Met Val Asp Arg Leu Pro Thr Val Leu Phe
            100                 105                 110
Phe Pro Cys Asn
        115
```